

# NORDISKE STUDIER I LEKSIKOGRAFI

Rapport fra Konferanse om leksikografi i Norden  
28.-31. mai 1991



# NORDISKE STUDIER I LEKSIKOGRAFI

Rapport fra Konferanse om leksikografi i Norden  
28.-31.- mai 1991

Redigert av Ruth Vatvedt Fjeld

SKRIFTER UTGITT AV NORDISK FORENING FOR LEKSIKOGRAFI  
SKRIFT NR. 1

© Nordisk forening for leksikografi. 1992

Trykk: Falch Hurtigtrykk as, Oslo

ISBN 82-91121-00-1

## Forord

I dagene 28-31 mai 1991 ble Konferanse om leksikografi i Norden holdt i Oslo. Denne rapporten inneholder de aller fleste foredragene som ble holdt på konferansen. Foredragene er bearbeidet til artikler, og relevante momenter som kom fram i diskusjonene, er innarbeidet. Konferansen ble arrangert av Avdeling for leksikografi ved Institutt for nordistikk og allmenn litteraturvitenskap, Universitetet i Oslo, i samarbeid med Nordisk språksekretariat. Konferansekomiteen bestod av

Ruth Vatvedt Fjeld, Avdeling for leksikografi  
Dag Gundersen, " " "  
Arnbjörg Hageberg, " " "  
Magnus Rindal, " " "  
Håvard Hjulstad, Rådet for teknisk terminologi  
Ståle Løland, Nordisk språksekretariat

Den samme komiteen har fungert som redaksjonskomite. Det redaksjonelle arbeidet har vært utført av Ruth Vatvedt Fjeld i samarbeid med Anne Torunn Engø ved Avdeling for leksikografi. Arne Jørgensen har vært rådgiver for den datatekniske delen av redaksjonsarbeidet. Medlemmene i redaksjonskomiteen har lest hver sin del av artiklene i rapporten, og har sammen med redaktøren vært med på å kommentere og godkjenne dem.

Konferansen fikk støtte fra Clara Lachmanns Fond, Letterstedtska Föreningen, Norges allmennvitenskapelige forskningsråd, Kirke-, utdannings- og forskningsdepartementet, Det historisk-filosofiske fakultet ved Universitetet i Oslo, og fra arrangørene Nordisk språksekretariat og Avdeling for leksikografi. Denne støtten har også gjort det mulig å utgi konferanserapporten.

På konferansen i mai 1991 ble Nordisk forening for leksikografi stiftet. Rapporten gis ut som det første skriftet i en serie som skal utgis i foreningens regi.

Oslo, mai 1992

Ruth Vatvedt Fjeld

## Nordisk forening for leksikografi

Ved avslutningen av den nordiske konferansen i leksikografi i Oslo 31.5.1991 ble det holdt stiftelsesmøte for Nordisk forening for leksikografi. Foreningen håper å samle flest mulig medlemmer blant alle som er virksomme - eller interesserte - i nordisk leksikografi. Ved slutten av sitt første år har foreningen ca. 80 medlemmer. Kontingenten for medlemskap i foreningen er for tiden nkr 150.

Foreningen planlegger å starte et tidsskrift og en skriftserie. Første publikasjon i skriftserien er denne konferanserapporten. Den blir utdelt gratis til foreningens medlemmer. Så langt opplaget rekker, gjelder dette også for dem som melder seg inn fra nå av.

Neste publikasjon i skriftserien er tenkt å bli en oversikt over nålevende utøvere av nordisk leksikografi, som i fremtiden kan oppdateres med passende mellomrom. Foreningens sekretariat sender ut spørreskjema til så vel medlemmer som ikke-medlemmer som vi kommer i kontakt med eller som tar kontakt med oss. Vårt sekretariat er

Nordisk språksekretariat  
Postboks 8107 Dep  
N-0032 Oslo  
Tlf. (02) 42 05 70  
Telefaks: (02) 42 76 76  
Foreningens postgironummer i Norge: 0824 0283116

Foreningens første styre består av følgende:

Anna Braasch (Danmark)  
Dag Gundersen (Norge, leder)  
Jón Hilmar Jónsson (Island)  
Nina Martola (Finland)  
Bo Svensén (Sverige)

### Vedtekter for Nordisk forening for leksikografi

§ 1 Foreningens navn er Nordisk forening for leksikografi

§ 2 Foreningens formål er å arbeide for å utvikle praktisk og vitenskapelig ordboksarbeid i Norden, og å fremme nordisk samarbeid om leksikografi.

§ 3 Alle som er interessert i leksikografiske spørsmål, kan bli medlem av foreningen. Medlemmene betaler en årlig kontingent, som fastsettes av generalforsamlingen. Medlemmene får foreningens publikasjoner, invitasjoner til arrangementer og rett til å delta i organisasjonsarbeidet.

§ 4 Generalforsamlingen velger et styre, som består av en leder, fire styremedlemmer og to varamedlemmer. Det skal tas nordiske hensyn ved sammensetningen av styret. Lederen og varamedlemmene velges for en periode på to år. Resten av styret velges for en periode på fire år, men slik at lederen og to styremedlemmer står på valg ved hver generalforsamling. Styret

konstituerer seg selv. I tillegg til lederen består styret av en nestleder, en sekretær, en kasserer, ett styremedlem og to varamedlemmer. Styremedlemmene kan velges høyst tre ganger.

Styret møtes minst en gang om året. Styret skal drive foreningen innenfor rammen av formålsparagrafen og de vedtak som generalforsamlingen gjør.

§ 5 Nordisk forening for leksikografi samarbeider med Nordisk språksekretariat om foreningens sekretariatsfunksjoner.

§ 6 Ordinær generalforsamling holdes annethvert år, fortrinnsvis i tilknytning til et faglig arrangement. Generalforsamlingen innkalles med en måneds varsel. Forslag til generalforsamlingen skal være styret i hende minst seks måneder før generalforsamlingen holdes. Styret sørger for at forslag utsendes sammen med innkallingen til generalforsamlingen. Saklista skal inneholde minst disse punktene:

- 1 Valg av ordstyrer
- 2 Beretning om virksomheten
- 3 Revidert regnskap
- 4 Fastsetting av kontingent
- 5 Valg av leder, to styremedlemmer og to varamedlemmer
- 6 Valg av valgkomité
- 7 Valg av revisor og revisorsuppleant
- 8 Forslag
- 9 Faglige arrangementer
- 10 Tid og sted for neste generalforsamling
- 11 Eventuelt

§ 7 Regnskapsåret er kalenderåret. Foreningen tegnes av leder og kasserer i fellesskap.

§ 8 Ekstraordinær generalforsamling holdes når styret beslutter det, eller når 1/4 av medlemmene skriftlig krever det. Ekstraordinær generalforsamling innkalles senest en måned og holdes innen to måneder etter at det har kommet krav om det.

§ 9 Forslag til endringer av disse vedtektene skal være styret i hende minst seks måneder før generalforsamlingen. Endringer av disse vedtektene skal vedtas på en generalforsamling med 2/3 flertall av de fremmøtte.

§ 10 Forslag om å oppløse foreningen skal være styret i hende minst seks måneder før generalforsamlingen. Forslaget skal vedtas på to generalforsamlinger med 2/3 flertall. Hvis foreningen blir oppløst, skal det avgjøres med vanlig flertall hva som skal skje med foreningens midler.

Inge Lønning

## Ordets makt

Hilsningstale ved åpningen av nordisk leksikografikonferanse  
Universitetet i Oslo 28. mai 1991

Kjære leksikografer  
fra alle de nordiske land!

Det er et privilegium å kunne ønske dere velkommen til det overdådige norske forår, til hovedstaden, til hovedstadens universitet og til Blindern. Så vidt jeg har forstått, er dette den første nordiske kongress for fagområdet leksikografi. Det gir begivenheten en særlig rang, og gjør det desto mer kjærkomment at dere har valgt å la den gå av stabelen her hos oss, ved Norges eldste og største universitet. Jeg håper at dere vil finne arbeidsmiljøet brukbart både innendørs og utendørs, og at det faglige og kollegiale utbytte må bli så godt at kongressen for fremtiden blir en selvsagt og uunnværlig fellestradisjon for de leksikografiske fagmiljøene i Norden.

Om det nordiske samarbeidet kan mangt sies, som vi alle vet. Især når en ny sesjon i Nordisk Råd står for døren, har kommentarene en tendens til å bli lett overbærende i tonen; de store politiske resultater er det ikke alltid like lett å få øye på. Ikke desto mindre vet vi jo godt at vi er begunstiget, vi som bebod det felles nordiske kulturrom. Vi er begunstiget med et språklig fellesskap som gjør at landegrensene ikke representerer kommunikasjonsbarrierer av betydning. Og vi er begunstiget med en noe nær ideell blanding av likheter og ulikheter oss imellom. De nordiske land er så like at sammenligninger vil være mulige og nærliggende å gjennomføre på de aller fleste områder, og samtidig så ulike at sammenligningene vil være interessante og bidra til å kaste lys over fellestrekk såvel som over egenarten i det enkelte folks kulturtradisjon.

Det siste tiår av det 2. årtusen er internasjonaliseringens tiår. Ingen som er knyttet til en institusjon for forskning og høyere utdanning kan unngå å ta opp den utfordring som ligger i at jernteppet er falt, og at de ideologiske kommunikasjonssperrene i Europa er borte. Det forekommer rett som det er - så i norsk debatt i hvert fall - at det nordiske og det europeiske stilles opp mot hverandre som om det skulle dreie seg om et enten eller. Enten nordisk samarbeide eller europeisk integrasjon, - det vil si: i den utstrekning det europeiske fellesskap utvides til å omfatte alle eller et flertall av de nordiske land, vil det nordiske samarbeide miste sin interesse.

En slik tankegang må bero på en åpenbar misforståelse. For førti år siden kunne det kanskje ha et snev av mening å betrakte et nordisk forsvarsforbund som alternativ til NATO-medlemskap. I dag har det ingen mening å se på nordisk samarbeid som et alternativ til europeisk integrasjon. Derimot har det særdeles god mening å se et styrket, utvidet og fordypet nordisk samarbeide som en uunnværlig bestanddel i den internasjonaliseringsprosess vi er på vei inn i. Jo viktigere det globale og det europeiske perspektiv blir, desto viktigere blir også vårt nordiske samarbeide. Å kommunisere - det betyr som kjent det enkleste og



vanskeligste av alt: å dele - og etablere gjensidig forståelse lærer man nå en gang bare ved å gå veien gjennom det nære, naboforholdet, til det fjernere.

Å være leksikograf og forvalte ord i disse tider må være en misunnelsesverdig oppgave. Den situasjon dere befinner dere i, kaller faktisk på den type epokeperspektiv som vi i min profesjonelle språkverden, den teologiske, ville være tilbøyelige til å bringe til uttrykk ved å minne om viktigheten av å kjenne sin besøkelsestid. Den kommunikasjonsoppgave vi står overfor i 1990-tallets Europa, vil aldri kunne løses gjennom språklig og kulturell standardisering, eliminering av individuell /nasjonal egenart, reduksjon av mangfold. Den vil bare kunne løses gjennom bevisst dyrkelse av mangfoldet, individualiteten, egenarten som vilkår for et fellesskap tuftet på likeverd og gjensidig respekt.

Å oversette oppfattes ofte som en mekanisk leksikalsk prosess, nærmest som et stykke enkelt håndverksarbeid. Den som har forsøkt å flytte et budskap fra én tidsalder til en annen, fra et språk til et annet, fra et meningsunivers til et annet, vil vite at ingenting kan være mer feilaktig. Vår menneskelige kommunikasjon er en altfor sammensatt og mangedimensjonal størrelse til at den lar seg demontere og sette sammen igjen ved hjelp av noen form for mekaniske øvelser.

Hvor vanskelig, og derfor også hvor utrolig spennende oversetteroppgaven er, gikk for alvor opp for meg da jeg for en del år siden arbeidet med en norsk Lutherutgave. Å overføre Luthers 1500-talls tysk og latin til et lesbart 1980-talls norsk kostet et betydelig antall nattetimer, men gav innsikt og forståelse av tekstene som aldri kunne ha vært mulig å vinne på annen måte. Og til belysning av den særegne forståelsesprosess som en oversettelse er finnes det fortsatt visdom å hente i Luthers egen refleksjon omkring kunsten å oversette. I "Sendbrief vom Dolmetschen" sier han - som svar til kritikken mot hans tyske oversettelse av Det Nye testamente (1522) - at det nytter lite å spørre bokstavene i det latinske språk hvorledes man skal snakke tysk. Den som skal oversette, må gå på torvet og lese på munnen til den alminnelige kvinne og mann, så kan han gå tilbake til sitt skrivebord og sette over budskapet fra den latinske eller greske tekst til tysk.

Det er et stykke oversettervisdom som knapt har mistet noe av sin aktualitet. Ikke er det noen grunn til å tro at den kommer til å miste den heller. For om vi aldri så mye lever i billedmedienes tidsalder, er det fortsatt slik at vi mennesker er språklige vesener, henvist til å løse våre samlivsproblemer - i alle sammenhenger og i alle betydninger av ordet - ved ordets hjelp. Ordets makt blir ikke mindre om bildene og billedskjermene øker aldri så mye i omfang. Derfor forblir det en livs-viktig kulturoppgave å ta vare på ordene og granske deres fortid, nåtid og fremtid. For det er ordene som binder generasjonene sammen og gir oss historie.

Kjære nordiske leksikografer, - dere behøver ikke å frykte for arbeidsløshet! Lykke til med den første i en forhåpentligvis lang rekke av nordiske kongresser for et livsviktig fagområde.



# Innhold

Ruth Vatvedt Fjeld: <i>Innledning</i>	s 6
Del 1:	
Bo Svensén: <i>Lexikografi och lexikografiska produkter — några grundbegrepp</i>	s 9
Bo Ralph: <i>Ord på väg. Några reflexioner kring nordisk lexikografi</i>	s 28
Del 2:	
Kristina Nikula: <i>Deskriptiva ordböcker - finns dom?</i>	s 43
Marit Hovdenak: <i>Norske skuleordliste</i>	s 53
Jørn-Otto Akø: <i>Gråsoner i norske ordbøker</i>	s 65
Dag Gundersen: <i>Om leksikografi som undervisningsemne</i>	s 76
Jón Hilmar Jónsson: <i>Fra en passiv til en aktiv ordbok. Det kombinatoriske aspektet i fokus</i>	s 88
Grete Duvå, Anna-Lise Laursen, Lisbet Maidahl: <i>Brugerundersøgelse vedrørende oversættelse af fagtekst</i>	s 105
Del 3:	
Sven-Göran Malmgren: <i>Om behandlingen av polysema ord i enspråkiga ordböcker</i>	s 137
Peter Widell: <i>Om definitioner i almensproglige ordbøger</i>	s 144
Åke Viberg: <i>Tvårspråklig lexicologi med svenskan i centrum</i>	s 156
Sven Lange: <i>En databaserad begreppsordbok för den äldre svenskan</i>	s 166
Jens Erik Mogensen: <i>Varians og distinktivitet</i>	s 176
Rolf Gavare: <i>Lexikografiska principer för alfabetisk filering med dator</i>	s 184
Botolv Helleland: <i>Proprium og leksikografi</i>	s 190
Tanneke Schoonheim: <i>The description of names in the Vroegmiddelnederlands Woordenboek</i>	s 200
Ruth Vatvedt Fjeld: <i>Om ord og termer i juridisk språk</i>	s 210
Petter Henriksen: <i>Forholdet mellom encyklopedologi og leksikografi</i>	s 219
Del 4:	
Hans Jonsson: <i>De stora historiska ordböckernas särart</i>	s 231
Arnbjörg Hageberg: <i>Norsk Ordbok - i 1950 eit nybrottsverk, i 1991 ein anakronisme?</i>	s 235
Henning Bergenholtz: <i>Almensproglige informationer i fagordbøger</i>	s 244
Hans Kristian Mikkelsen: <i>Brikker til en typologi over danske fagordbøger</i>	s 260
Sandro Nielsen: <i>Forholdet mellem alordbøger og enkeltfagsordbøger</i>	s 275
Sonja Poulsen: <i>Kollokationsundersøgelse inden for engelsk merkantilt og økonomisk sprog med henblik på udarbejdelse af ordbogskoncept</i>	s 288
Christian Becker-Christensen: <i>Om opstilling af synonymer i synonym- og betydningsordbøger</i>	s 296
Anne-Line Graedler: <i>Orientering om prosjektet "Norsk anglisismeordbok"</i>	s 307
Erik Simensen: <i>Om ein plan for ei mellomnorsk ordbok — ordbok over skriftmålet i Noreg 1350-1550</i>	s 315
Viggo Sørensen: <i>Jysk Ordbog - en præsentation</i>	s 324

Tor Erik Jenstad: <i>Presentasjon av prosjektet "Ordbok over trøndermåla"</i>	s 339
Martin Gellerstam, Sven-Göran Malmgren og Maria Toporowska-Gronostaj: <i>Inlärningslexikon</i>	s 347
Helge Niska og Jorunn Frøili: <i>Tolkordlistor på invandrarpråk</i>	s 357
Del 5:	
Lars Svensson: <i>Om SAOB:s material och en bibliografisk databas</i>	s 369
Guðrún Kvaran: <i>En ordbog og dens informanter</i>	s 380
Helle Degnbol, Guðrún Ása Grímsdóttir, Bent Chr. Jacobsen, Jette Knudsen, Eva Rode & Christopher Sanders: <i>En ordbog er en database</i>	s 385
Ásta Svavarsdóttir, Jón Hilmar Jónsson, Kristín Bjarnadóttir: <i>Fra seddel- samlng til database: Leksikografisk analyse af islandske verber</i>	s 390
Christian-Emil Ore: <i>Dokumentasjonsprosjektet ved Det historisk-filosofiske fakultet, Universitetet i Oslo</i>	s 403
Jan Engh: <i>Leksikografi i IBM Norge</i>	s 409
Jostein Baustad: <i>Automatisk analyse av maskinleselige ordbøker til bruk i en orddatabase</i>	s 423
Valerij Berkov: <i>"Vulgært" i tospråklige ordbøker</i>	s 432
Hannele Seppinen: <i>Finska för utlänningar - uppgifter om ordböjning i den finsk-norska ordboken</i>	s 436
Rik Schutz: <i>Redigeringsprogrammet EDI</i>	s 441
Ole Norling-Christensen: <i>Struktureret redigering af ordbøger</i>	s 447
Merete K. Jørgensen: <i>Ordbogsartikler som skriver sig selv - ? Stednavne som kilde i Gammeldansk Ordbog</i>	s 455
Del 6:	
Göran Karlsson: <i>Empiriska synpunkter på tvåspråkig lexikografi</i>	s 465
Hans de Groot: <i>Från svensk-engelsk till svensk-nederländsk</i>	s 473
Rita Lenstrup: <i>Den selektive specialordbog</i>	s 478
Frede Boje og Anna Braasch: <i>Hvad får man skudt i skoene? Flerordsenheder i aktive ordbøger for mennesker og maskiner</i>	s 483
Patrick Nigel Chaffey: <i>Problemer i forbindelse med tospråklig juridisk leksikografi</i>	s 501
Eli Ellingsve: <i>Tvetydighet som kommunikasjonsproblem Betydningen av en utvetydig terminologi i internasjonal kommunikasjon</i>	s 509
Nina Martola: <i>Arbetet med Finsk-svensk ordbok: Normeringsproblem visavi finlandssvenskan</i>	s 515
Sven Tarp: <i>Encyklopædiske oplysninger i tosprogede tekniske ordbøger</i>	s 522
Tove Jacobsen: <i>En type bevægelsesverb i norsk, og hvordan de er behandlet i noen norsk-franske ordbøker</i>	s 532
Jan Hoel: <i>Norsk EF-terminologi. Om det pågående arbeidet med å registrere norsk terminologi knyttet til oversettelsen av deler av EFs regelverk til norsk</i>	s 540
Stikkordliste	s 549
Refererte ordbøker	s 566
Oversikt over forfatterne	s 559

Ruth Vatvedt Fjeld

## Innledning

Leksikografi er et fag med lange tradisjoner. Det å samle og systematisere ord, og definere dem eller sammenstille dem med ekvivalenter i andre språk, har vært gjort i over tusen år. Ordbøker har først og fremst en klart praktisk funksjon. Ved hjelp av dem kan språkbrukerne skaffe seg mer kunnskap om sitt eget språk, eller få hjelp til å lære eller forstå et språk de ikke kan. I en tid med stadig mer internasjonal kontakt er behovet for ordbøker åpenbart, noe som ofte blir påpekt i internasjonale organer, og blir prioritert i samarbeidsprogrammer. På den annen side er leksikografisk arbeid viktig for dokumentasjon av den faktiske språkbruken, og for å danne grunnlag for ny teoriutvikling om språk generelt. Grunnlaget for all leksikografi er systematisk observasjon av store mengder anvendt språk, og ofte er det få andre faginstusjoner eller fagområder som driver med slik systematisk observasjon. Leksikografiske arbeider er derfor viktige kilder for andre lingvistiske disipliner.

På grunn av de store og raske samfunnsendringene som skjer for tiden, både i politisk, demografisk og organisasjonsmessig forstand, er det bred enighet om at nordisk samarbeid bør styrkes. De nordiske språkene har mye til felles, mange vil regne flere av dem som varianter av samme språk. Dersom en skal komme vesentlige skritt videre i språkbeskrivelsen og kodifiseringen, er samarbeid en klar fordel, om ikke en forutsetning. Mer kontakt og bedre oversikt over det arbeidet som gjøres, kan føre til at en unngår dobbeltarbeid, og samarbeid kan gi inspirasjon og støtte der det er vanskelig å finne fram til gode metoder, eller der en savner teoretisk overbygning.

Det pågår nå en debatt i fagmiljøet om leksikografi er en egen vitenskap eller en ren ordinnsamling av ateoretisk karakter. For å vurdere dette forholdet er det nyttig å se på hvilke kriterier en kan legge til grunn for å bestemme begrepet vitenskap. Grunnleggende krav er at faget har sin egen teori og at det har sine egne vitenskapelige metoder. Inntil forholdsvis nylig har man sett på leksikografi som praktisk arbeid med utforming av ordbøker og ordlister, et rent samlings- og arkivarbeid. I den grad en har støttet seg på teori, har det vært regnet som allmenn lingvistisk teori, og metoden har ofte vært å avveie forskjellige praktiske hensyn mot hverandre, gjerne fra tilfelle til tilfelle, ja, så å si fra ord til ord. Helheten og konsekvensen i fremstillingen kan det da bli skralt med. Det en har hatt av leksikografisk teori, har vært nedfelt i forordene til ordbøkene, de samme forordene som er så berømt for aldri å bli lest. Årsaken til at de ikke leses, ligger kanskje i de dobbelte siktemål de ofte har hatt: Å informere brukerne om hvordan de best kan utnytte ordboka, og å gjøre rede for hvilke prinsipper som ligger til grunn for utvelgelse og systematisering.

Leksikografien har altså ikke hatt noen egen teori eller i hvert fall ikke en kodifisert teori. Leksikografi har vært praktisert etter hva en har lært av forgjengere i faget, som tidligere en far lærte sin sønn et håndverk. I nordisk sammenheng har det i enspråksleksikografien vært satt i gang store ordboksprosjekter som skulle omfatte hele ordmaterialet i det aktuelle språket, og et møysommelig og nitid innsamlingsarbeid har pågått, for noen i over hundre år. Når en kommer med i et slikt omfattende arbeid, der

redaksjonsprinsipper er satt ned eller har utviklet seg kanskje for generasjoner siden, er det naturlig at det daglige, konkrete redigeringsarbeidet kommer i forgrunnen, og at det blir satt av mindre tid til prinsipielle og teoretiske drøftinger. Den tospråklige leksikografien har stort sett levd en skjult tilværelse, og tospråklige ordbøker har vært produsert av enkeltpersoner på oppdrag fra forlag, altså mest som en rent kommersiell foreteelse, og da selvsagt styrt av økonomiske og kommersielle hensyn.

Først for ca 20 år siden kom det fart i teoridannelsen og det som i dag vanligvis kalles metaleksikografien, det vil si beskrivelser av ordbøker, deres funksjon i samfunnet og "oppskrifter" på hvordan de skal eller kan lages ut fra forskjellige målsettinger med ordboka.

Parallelt med utdanningsekspløsjonen og utviklingen av det skriftspråksavhengige samfunn skjedde det en nyorientering innen allmenn lingvistik. I strukturalismens glansdager på 60-tallet var semantikens rolle blitt sterkt nedtonet, og med den leksikon eller leksikologien. I Chomskys generative grammatikkmodell var semantikken stemoderlig behandlet, og leksikon ble fremstilt som en fri komponent som virket på de ferdige strukturene som var blitt generert av de andre grammatiske komponentene fra en "dypstruktur" til den rette "overflatestrukturen". Den siste operasjonen, innsetting av leksikon, ble lite utfyllende beskrevet.

Et første vendepunkt kom egentlig med Katz og Fodors semantiske teori i midten av 60-årene, men først mot slutten av 70-årene fikk semantikken en skikkelig plass i språkbeskrivelsene, og utover 80-tallet har leksikografi blitt viet stor oppmerksomhet i allmenn lingvistisk teori, f.eks. Kaplan & Bresnan 1982, Gazdar & al 1985 og Kay 1985. Leksikon har fått en fremtredende plass i modeller som LFG (lexical-functional grammar) og GPSG (generalized phrase structure grammar) og unifikasjonsgrammatikken, og utgjør der en basiskomponent i språkbeskrivelsen. Det har igjen ført til økt interesse for emner som tidligere har vært leksikografenes hovedanliggende, som morfologi og orddanningsregler. Denne teoriutviklingen kommer nå leksikografien til gode.

Datamaskinell språkbehandling har skutt fart i den samme perioden. Det gjelder både maskinell analyse av menneskeprodusert språk og maskinenes evne til å produsere menneskelignende språk ved hjelp av regler og leksikon. Det har gitt styrket håp om at det skal kunne gå an å oversette språk automatisk. Leksikografi har blitt et "nødvendig onde" i maskinoversettelse, dvs. at kunnskap om ords bruk og ords betydning er nødvendig for å kunne lage både regelbaserte og leksikonbaserte automatiske oversettelsesprogram. Rent praktisk har tilgjengelighet til leksikografisk kunnskap vært å hente hos leksikografene i deres arkiver og seddelsamlinger, og det har ført til en enorm oppvurdering av samlingenes verdi, men også til en kritisk gjennomgang av deres struktur og innhold.

Datamaskinell lagring og gjenfinning har gjort det lettere å systematisere store ordbaser og å gjøre korpusundersøkelser. Her har leksikografi og terminologi fått mye felles, og leksikografien har kunnet nyte godt av terminologisk metode.

Med oppbyggingen av det nye utdanningssystemet har universitetene fått konkurranse fra andre institusjoner som gir høyere, profesjonsrettet utdanning. Særlig har handelshøgskolene utviklet gode språkfaglige miljøer, der fremmedspråkundervisning og oversettelse har en fremtredende plass. Det har ført til at de tospråklige ordbøkene har fått mer oppmerksomhet og er blitt underlagt vitenskapelig og systematisk tenkning. Denne utviklingen har igjen virket inspirerende på de etablerte leksikografiske miljøene, der man mest har arbeidet med langsiktige, enspråklige prosjekter. Samtidig har oversettelse blitt utviklet til en egen fagdisiplin, oversettervitenskap, og det medfører selvfølgelig et økt behov for tospråklige ordbøker.

Ordbøkernes verdi som hjelpemiddel i menneskelig kommunikasjon er blitt oppvurdert etter hvert som stadig flere får en høyere utdanning og blir mer og mer avhengig av å kunne uttrykke seg skriftlig. At leksikografi i dag anses som en egen fagdisiplin, bekreftes av det faktum at det skrives teoretiske arbeider om leksikografi, at det gis undervisning på universitetsnivå i faget, at det skrives faghistorie, og at det organiseres konferanser og andre faglige fora (tidsskrifter, foreninger) for fagintern kommunikasjon. Ved Universitetet i Oslo ble det i 1989 slått fast i en komitéutredning at visse typer leksikografisk arbeid er vitenskapelig arbeid, og at det har samme meritterende kraft som annen forskning.

Et annet spørsmål er det om ordboksforfatterne vil eller kan anvende den teorien som er tilgjengelig. Her kommer ofte de tidsmessige og økonomiske sidene ved ordboksarbeidet inn, samt leksikografenes kompetansenivå. Særlig for den kommersielle leksikografien er det et vurderingsspørsmål. Men økt kunnskap om metaleksikografi vil gi bedre og sannsynligvis billigere ordbøker. Særlig er utviklingen av dataleksikografi - eller foreløpig mest datastøtta leksikografi - med på å gjøre produksjonen av ordbøker enklere og billigere.

Wiegand (1989) tar i en artikkel opp spørsmålet om leksikografi er en vitenskapelig disiplin. Han skiller da, slik det er vanlig å gjøre, mellom praktisk leksikografisk arbeid og studiet av disse produktene, dvs ordboksforskning eller metaleksikografi. Det praktiske innsamlings- og redigeringsarbeidet går han ut fra som en ikke-vitenskapelig disiplin (det forhindrer ikke at det må kunne betraktes som akademisk arbeid), det er ordboksforskningen eller metaleksikografien han drøfter, også kalt språkleksikografi (dvs at ordbøkene gir leserne kunnskap om språket), til skille fra terminologi og encyklopedi eller sakleksikografi. Det er ofte en terminologisk forvirring her. Wiegand tar utgangspunkt i Posner (1987 og 1988), som har satt opp tre avdelinger eller mengder med kriterier som må være oppfylt for at noe skal kunne regnes som vitenskapelig disiplin. Den første mengden inneholder fem kriterier som må oppfylles; nemlig at faget har et fast gjenstandsområde, en bestemt mengde perspektiver (bestemt ut fra studiets mål), bestemte metoder (dvs at fremgangsmåter kan gjentas og kan læres bort), et bestemt vitenskapskorpus og bestemte fremstillingsmidler. Mengde 2 skjerper inn de fem kravene, slik at det ikke er nok bare å oppfylle dem, det må også være en viss sammenheng mellom de aktivitetene som drives innen faget, dvs at det settes krav til homogenitet og dominans. Hvis alle disse betingelsene er oppfylt, kan en regne aktiviteten som en faglig disiplin i vitenskapsteoretisk forstand. (Wiegand definerer vitenskap som teoridannende forskning.) I mengde 3 stilles det videre som krav at faget må ha en viss totalitet eller dekningsgrad innen feltet, og at det må gi uttrykk for verdinøytralitet for at det skal kunne betraktes som en vitenskapelig disiplin. Disse kriteriene bruker så Wiegand i en vurdering av metaleksikografi. Han konkluderer med at den ikke entydig er en vitenskapelig disiplin, men at den er i ferd med å utvikle seg til å bli det. Hovedankepunktet hans er at det ikke er utviklet en konsistent leksikografisk teori. Det virker som om det settes svært strenge krav til at noe skal kunne kalles en vitenskapelig disiplin, og det kan hende at det ikke fins noen disipliner som oppfyller alle disse kravene, de får heller betraktes som idealiserte forestillinger om vitenskapen. Og man kan i hvert fall ikke lenger hevde at leksikografi er et teoriløst fag, dvs at det ikke fins metaleksikografisk litteratur tilgjengelig. I 1989 kom første bind i det store verket *Wörterbücher/Dictionaries/Dictionnaires. Ein internationales Handbuch zur Lexikographie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie*), redigert av Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand og Ladislav Zgusta (HSK.5). Det er femte bind i gigantserien *Handbücher zur Sprach- und Kommunikationswissenschaft* og er et prosjekt som tar mål av seg til å beskrive alle sider ved hele verdens ordbokssamlinger og ordboksproduksjon, en slags leksikografiens tesaurus. Totalt har verket 3355 sider, det vil si netto ca 3000 sider teoretisk utgreiing om

faget. Det viser den store interesse det er for teoretisk leksikografi i fagmiljøene, men også at det er store luker innen faget, både med hensyn til teoridanning og til forskjellige ordbokstyper. Når en har foran seg dette gigantverket, kan en fristes til å sitere hva en av de norske store ordboksmenn, Ivar Aasen, sa da han viste til sine tyske forgjengere i faget og roste dem for deres grundighet: "De tydske Sprogmand have i denne Sag givet et lysende eksempel" (Fortale til *Norsk Ordbog* 1850:IX).

I forbindelse med rapporten fra Konferanse om leksikografi i Norden tjener HSK.5 som mer enn et lysende eksempel. I den grad det har latt seg gjøre ved redigeringen av artiklene, er det forsøkt å følge samme ordningsprinsipper som der. For selv om konferansen ble større enn det på forhånd var antatt, selv om deltakelsen med foredrag var overveldende, og selv om oppfølgingen med å omarbeide foredragene til artikler som kunne presenteres i en rapport, har vært over all forventning, så ligger det klare begrensninger i emneomfanget. For det første er omfanget begrenset til nordisk leksikografi. Nordisk leksikografi er her definert som leksikografisk og metaleksikografisk arbeid om emner som på en eller annen måte er relatert til et av de nordiske språkene. Videre har det ikke vært innkalt foredragsholdere i forhold til et planlagt program eller en disposisjon, alle har meldt seg på med emner som de er opptatt av eller arbeider med for tiden. Derfor inneholder rapporten mer et øyeblikksbilde av aktiviteten i nordisk leksikografi våren 1991 enn en oversikt over nordisk leksikografi. Men å sammenholde dette øyeblikksbildet med det internasjonale oversiktsverket kan gi et inntrykk av tilstanden sett i relasjon til hva som kanskje burde eller kunne være gjort innenfor faget i nordisk sammenheng. For det andre ble det lagt ned et stort arbeid i organiseringen av HSK.5, og det er et ønske om å dra nytte av dette arbeidet ved å organisere konferanserapporten i samsvar med den. For det tredje er HSK.5 det verket som uten konkurranse er oftest referert til i de artiklene som rapporten inneholder. Den fungerer altså som et referanseverk for nordiske leksikografer flest. Jeg vil likevel sitere og slutte meg til følgende fra forordet til redaktørene:

Die Erarbeitung einer bestimmten Gliederung führt stets zugleich zu einer bestimmten Perspektive, so dass zwangsläufig eine unterschiedliche Gewichtung von Aspekten des Gegenstandsbereiches erfolgt.  
(s VIII).

Konferanserapporten er delt opp i seks deler. Den første delen bryter med organiseringen av HSK.5. Artiklene presenterer grunnleggende problemstillinger og definisjoner, gir en oversikt over forskjellige ordbokstyper (Svensén), og gir en plassering av leksikografien i tid og rom (Ralph). Disse er artikkelversjonene av innlednings- og avslutningsforedraget på konferansen, og vil gi god bakgrunn for den videre lesningen.

Del 2 tar opp sider ved forholdet mellom leksikografi og samfunn og hvordan leksikografer kan være med og styre et språks utvikling (Nikula, Hovdenak, Akø), undervisning i leksikografi (Gundersen) og hvordan ordbøkene må være for at forskjellige typer brukere skal ha best mulig utbytte av dem (Maidahl m fl, Jonsson).

Del 3 tar opp teori i enspråksleksikografien. Først står de artiklene som på en eller annen måte tar opp de mer generelle teoretiske problemene i allmennordbøker; hvordan generell semantisk teori kan anvendes i leksikografien (Malmgren, Widell, Viberg, Lange, Mogensen), og hvordan ordbøkene kan organiseres eller struktureres (Gavare). Deretter er det samlet artikler som tar opp enkeltproblemer innen allmennspråksleksikografien, særlig utvalgsriterier for lemmaer i allmennspråklige ordbøker (Henriksen, Helleland, Schoonheim, Fjeld).



Del 4 består av artikler med drøftinger om ordbokstypologi. De fem første artiklene tar for seg generelle ordbokstypologiske problemer (Jonsson, Hageberg, Bergenholzt, Mikkelsen, Nielsen), deretter kommer tanker rundt spesielle typer ordbøker ordnet etter syntagmatiske (Poulsen) og paradigmatiske (Becker-Christensen) prinsipper. Andre spesialordbøker har sitt grunnlag i avvik fra et tenkt standardspråk: lånnord (Graedler), ord fra en bestemt historisk periode (Simensen), fra bestemte geografiske områder (Sørensen, Jenstad) og for å dekke spesielle brukerbehov (Gellerstam m fl, Niska & Frøili).

Del 5 inneholder artikler som behandler emner om leksikografisk metode, hvordan en finner fram til det stoffet en vil eller bør ha med (Svensson, Kváran), hvordan en kan systematisere og oppbevare det en har samlet, bl a ved hjelp av datateknologi (Degnbol m fl, Svavarsdottir m fl, Ore, Engh, Baustad), hvilke lingvistiske etiketter oppslagsordene skal forsynes med (Berkov, Seppinen), og hvordan det praktiske redigeringsarbeidet kan utføres (Schutz, Norling-Christensen, Jørgensen). Del 5 er en av de fyldigste i rapporten. Det er naturlig, siden det skjer en veldig metodeutvikling i faget for tiden. Ny teknologi og nye teorier krever nye metoder.

Den siste delen, del 6, omhandler tospråklig og flerspråklig leksikografi, både allmenne problemer (Karlsson, de Groot), utvalgsproblematikk (Lenstrup) og ekvivalensspørsmål (Chaffey, Boje & Braasch, Ellingsve, Martola, Tarp, Jacobsen). Den siste artikkelen (Hoel) presenterer en flerspråklig termbase, et prosjekt som ligger i ytterkanten av den egentlige leksikografien, men som sannsynligvis illustrerer arbeidsmetoder og -oppgaver som fremtidens leksikografer må forholde seg til.

Mange av artiklene tar opp flere aspekter ved leksikografisk arbeid, og ved å velge et annet perspektiv ved lesingen, kunne en fått en helt annen organisering av rapporten. Noen emner er behandlet av flere og danner klart sammenhengende avdelinger i rapporten, andre emner er mer alenestående, og mange er helt ubehandlet. Det kunne vært mer utfyllende behandling av emner om leksikografiens oppgaver og virkning i samfunnet, og hvilken status leksikografien har innen forskersamfunnet ellers, hvordan ordbøker brukes (eller kan brukes) i språklig planleggingsarbeid, eller om ordbøker som kommersielle produkter. Ordboksbruk og adaptering av ordbøker til visse brukergrupper har heller ingen fremtredende plass i rapporten, og det er ingen artikkel som tar opp leksikografiens historie i Norden. Helt generelle metaleksikografiske problemstillinger er også dårlig representert; selv om mange emner blir tatt opp i forskjellige artikler, har de ikke vært artiklenes hovedanliggende. Det hadde også vært nyttig å få et bilde av ordbokssituasjonen i Norden, hvilke språkpar det fins ordbøker for, og hvilke spesialordbøker som fins på de enkelte språkene. F eks har vi ikke noen egen barneordbok for norsk, mens det fins en for dansk og en for islandsk. Vi vet heller ikke hva slags konstruksjonsordbøker som fins for de forskjellige språkene, eller hvilke dialekter/geografiske områder som har egne ordbøker, hvilke frekvensordbøker som er utarbeidet, hvilke historiske perioder som er belagt i ordbøker, hvilke tekster/forfattere det er utarbeidet ordbøker på grunnlag av, osv. En grundig oversikt ville være til hjelp når en skal studere språk eller lage ordbøker, og en oversikt ville også være en inspirasjon for nye ordboksprosjekter.

For å hjelpe leserne med å finne fram til de emnene de søker, er det utarbeidet et fyldig stikkordregister, som både opplyser om hvilke termer som er anvendt, og om hvilke emner som har vært behandlet. I tillegg er det laget et eget register over de ordbøkene som er direkte nevnt i artiklene. Denne oversikten håper vi kan være nyttig som inngang til omtale av ordbøker en bruker eller velger å ha til forelegg for nye ordboksprosjekter.

Tradisjonen i praktisk leksikografi har langt på vei vært nettopp å bruke tidligere utgitte ordbøker som modell, ofte også som kilde, for arbeidet. I få fag er

akkumulasjonsteorien så godt illustrert som i leksikografien. Faget står i dag foran store og spennende utfordringer, med nye teorier, ny metode og ny teknologi for hånden. Det er å håpe at rapporten er med på å spre kunnskap om disse teoriene og metodene, og at man kan forene den praktiske innsikt og veldige språkkunnskap som forgjengerne i faget har gitt, med denne nye viten. Men det må ikke bli teoriutvikling for teoriens egen skyld, kunnskapen bør føre til at det skapes gode leksikografiske produkter i fremtiden, en fremtid der enkeltmenneskene virkelig trenger å kunne tilegne seg kunnskap om språk og språkbruk for å fungere godt i samfunnet.

### Litteratur

- Chomsky, Noam. 1957. *Syntactic Structures*. Haag
- Chomsky, Noam. 1986. *Knowledge of Language*. New York
- Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, Mass.
- Hausmann, F.J., Reichmann, O., Wiegand, H.E. & Zgusta, L. 1986-91. *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*. Berlin, New York
- Katz, J.J. & Fodor, J.A. 1963. "The structure of a semantic theory". *Language* 39:170-210
- Kaplan, Ronald & Bresnan, Joan. 1982. *Lexical-Functional Grammar: A Formal System for Grammatical Representation*. I: Bresnan (ed): *The Mental Representation of Grammatical Relations*. Cambridge, Mass.
- Kay, M. 1985. "Parsing in functional unification grammar". I: D.R. Dowty, L. Karttunen & A.M. Zwicky (eds): *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*. Cambridge, Mass.
- Posner, Roland. 1987. *The relationship between individual disciplines and interdisciplinary approaches*. Berlin
- Posner, Roland. 1988. "What is an Academic Discipline?" I: *Gedankenzeichen. Festschrift für Klaus Oehler zum 60. Geburtstag*. Tübingen
- Sells, Peter. 1985. *Lectures on contemporary syntactic theories*. Stanford
- Utvalget for forskningsspørsmål, Det historisk-filosofiske fakultet ved Universitetet i Oslo. 1989. *Vurdering av ordboks- og arkivarbeid i forhold til vitenskapelige kvalifikasjoner*. Upublisert
- Wiegand, H.E. 1989. "Der gegenwärtige Status der Lexikographie und ihr Verhältnis zu anderen Disziplinen. I: Hausman & al: *Wörterbücher / Dictionaries / Dictionnaires*. Berlin, New York
- Aasen, Ivar. 1850. *Norsk Ordbog*. Kristiania

Del 1



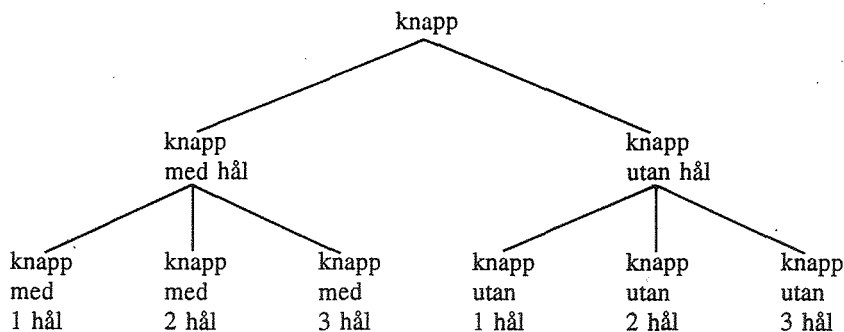
Bo Svensén

## Lexikografi och lexikografiska produkter. Några grundbegrepp

Syftet med detta bidrag är att rekapitulera några för ämnet lexikografi grundläggande begrepp genom att presentera olika sätt att beskriva och systematisera olika typer av ordböcker. De huvudaspekter som tas upp är dels ordbokens förhållande till världen, till användarna, till utgångsmaterialet och till det språkliga tecknet, dels antalet språk som behandlas i ordboken, de informationskategorier som förekommer i den samt dess struktur. Dessutom avgränsas den språkligt inriktade ordboken från två besläktade produkter: encyklopedin och den terminologiska ordlistan.

### O. Inledning

Strindberg skildrar i "De lycksaliges ö" en person som är expert på knappor och som genomför en vetenskaplig systematisering av knapparna. En del av det system han utarbetade kan åskådliggöras som i Figur 1:



Figur 1

Vederbörandes utredning av knapparnas systematik gav honom förstås en professur i ämnet, vars benämning, *knappologi*, fortfarande används i svenskan och enligt Svensk ordbok betyder "systematisering av värdelösa fakta (med vetenskapliga anspråk)".

När jag nu skall försöka rekapitulera några för ämnet lexikografi grundläggande begrepp genom att gå igenom olika sätt att beskriva och systematisera olika typer av ordböcker, hoppas jag kunna förekomma alla anklagelser för knappologi genom att påpeka

- a) att det inte rör sig om helt värdelösa fakta, eftersom den typologiska aspekten utgör en viktig del av metalexikografien, både teoretiskt och praktiskt, och
- b) att det jag kommer att säga inte har några vetenskapliga anspråk.

En beskrivning av en ordboks egenskaper kan exempelvis ta upp följande huvudaspekter:

- Ordbokens förhållande till världen.
- Ordbokens förhållande till det språkliga utgångsmaterialet.
- Ordbokens förhållande till det språkliga tecknet, nämligen vilken komponent av tecknet den tar sikte på
- Antalet språk som behandlas i ordboken: ett eller två (eller möjligen flera).
- Informationskategorier som förekommer i ordboken.
- Ordbokens struktur.
- Ordbokens förhållande till användarna.

### 1. Ordbokens förhållande till världen

Om vi börjar från början och ser efter hur ordboken förhåller sig till världen, kan vi konstatera att föremålet för beskrivningen i en ordbok normalt är språkliga enheter, inte företeelser i världen utanför språket. Häri skiljer sig ordboken från en i många avseenden starkt besläktad produkt, *encyklopedin* (uppslagsboken). Grovt uttryckt beskriver encyklopedin världen som helhet, medan ordboken bara ger information om det kommunikationssystem med vars hjälp människor utbyter meddelanden om världen.

Men här finns det förstås gränsfall:

- Det ena är den s.k. *encyklopediska ordboken*, som kan beskrivas som en språklig ordbok med systematiska encyklopediska inslag. Här finns det flera olika former, t.ex. encyklopediska ordböcker som inom artiklarna gör åtskillnad mellan språkligt och encyklopediskt och sådana som blandar språkligt och encyklopediskt<sup>1</sup>.
- Det omvända fallet förekommer också, alltså en *encyklopedi med språkbeskrivande inslag*. Detta kan innebära att man tar med "svåra ord" eller "främmande ord" och förklarar dem med synonymer eller liknande, utan att ge encyklopedisk information. Det är också vanligt att encyklopedier ger etymologiska upplysningar om uppslagsorden, vilket egentligen är främmande för syftet med denna typ av uppslagsverk.

Med det här har jag dock inte velat påstå att en språkligt inriktad ordbok inte kan eller får innehålla information om världen. Tvärtom måste den ofta appellera till användarens kunskaper om världen för att precisera den information som ges om språket.

Jag lämnar nu den här delen av ämnet och kommer i fortsättningen med termen "ordbok" att mena en helt och hållet språkligt inriktad produkt.

## 2. Ordbokens förhållande till det språkliga utgångsmaterialet

2.1 En ordbok kan beskrivas utifrån det utgångsmaterial (den *korpus*) som ligger till grund för den.

2.1.1 Korpus kan utgöras antingen av ett visst *språk* eller av en viss *text*. Typiska textrelaterade produkter är då *konkordanser*, *index* och liknande. Hit hör också *exegetiska ordböcker*, dvs. sådana som förklarar ord och uttryck i den ordning de uppträder i texten och vid förklarandet av ett visst ord eller uttryck bara anger den betydelse etc. som är aktuell just där.

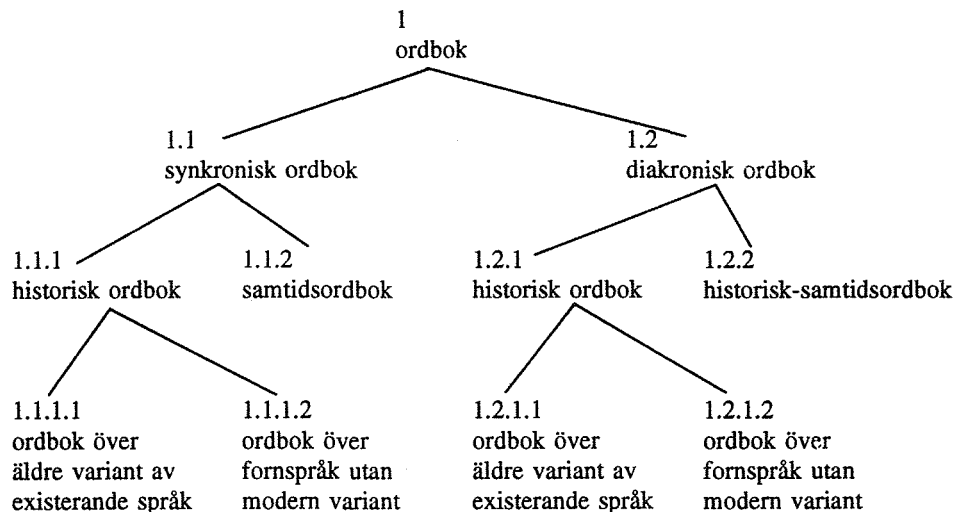
2.1.2 En annan distinktion kan göras mellan *öppen* och *sluten* korpus. En viss text utgör naturligtvis en sluten korpus. Men även språk och språkvarianter kan utgöra slutna korpusar: ett dött språk som den antika grekiskan kan i princip betraktas som en sluten korpus, även om det är fullt tänkbart att man i framtiden kan komma att återfinna något av de 113 dramer som vi vet att Sofokles skrev men som är försvunna.

2.1.3 Man kan också göra skillnad mellan en *kollektivspråklig* och en *individualspråklig* korpus, alltså mellan en språkgemenskaps språk och en författares språk.

2.2 En ordbok kan också beskrivas utifrån det sätt på vilket lexikografen har förhållit sig inför denna korpus. Han kan för det första välja antingen att behandla i stort sett alla förekomster av alla ord eller att arbeta selektivt. Det förra alternativet är i praktiken meningsfullt bara när korpus är sluten och har begränsad storlek, och i fortsättningen utgår jag från att han arbetar selektivt. Det urval han därvid gör kan beskrivas utifrån flera olika aspekter, beroende på vilket syfte ordboken har.

2.2.1 Den språkliga beskrivningen kan t.ex. vinklas efter ett visst *tidsperspektiv*, och då får man följande möjligheter.

Ordboken kan vara *synkronisk*, dvs. beskriva språkets tillstånd under ett begränsat skede, eller *diakronisk*, dvs. beskriva språkets utveckling under en längre period. En annan distinktion kan göras mellan *historiska ordböcker* och *samtidsordböcker*. Kombinerar vi dessa två indelningar får vi i ett system som i Figur 2.



Figur 2.

Bland de *synkroniska* historiska ordböckerna kan man tänka sig sådana som beskriver dels äldre varianter av fortfarande existerande språk, t.ex. en ordbok över Chaucer (1.1.1.1), dels fornspråk utan motsvarande modern variant, t.ex. en ordbok över de tre klassiska grekiska tragödem (1.1.1.2). De synkroniska samtidsordböckerna (1.1.2) är förstås de vanligaste, den typ lekmannen normalt associerar till när han hör termen "ordbok". Som en extrem variant av synkronisk samtidsordbok skulle man kunna beteckna *nyordboken*.

Men ingen synkronisk samtidsordbok för allmänt bruk är i praktiken helt synkronisk. Vid en given tidpunkt lever nämligen minst tre från språklig synpunkt relevanta användargenerationer med delvis olika språkbruk, som ordboken måste ta hänsyn till. Dessutom måste ordboken också förklara det språkbruk som förekommer i sådan äldre litteratur som betraktas som "levande" och fortfarande läses av dessa användargrupper<sup>2</sup>.

Bland de *diakroniska* historiska ordböckerna kan man tänka sig samma indelning som under de synkroniska. Motsvarigheten till *Chaucerordboken* skulle här bli en ordbok över hela den medelengelska språkperioden (1.2.1.1), och motsvarigheten till de grekiska tragödem skulle bli en ordbok över antik grekiska (1.2.1.2). Den sista kategorin, den historiska samtidsordboken (1.2.2), skulle då bli en ordbok som beskriver språkets utveckling under en längre period inklusive samtiden. Typiska exponenter är de stora nationella ordboksverken av typ *Svenska Akademiens ordbok*, *Oxford English Dictionary* osv. Ett specialfall i den här gruppen är den *etymologiska ordboken*.

**2.2.2** Det finns andra dimensioner än den tidsmässiga att ta hänsyn till. Ordboken kan exempelvis ta sikte på den *geografiska* aspekten och beskriva regionalt språk, varigenom vi får exempelvis en *dialektordbok*. Den kan också beskriva *varianter av ett visst språk i ett visst land*, t.ex. en ordbok över kanadensisk franska, över australiensisk engelska eller över schweizisk tyska. Den kan också välja att lista enbart ord av ett visst ursprung, t.ex. i en *ordbok över främmande ord*.



**2.2.3** Ordboken kan vinkla urvalet så att den enbart behandlar ord som tillhör *stilmässigt och socialt* betingade varianter av språket, och resultatet kan exempelvis bli en *slangordbok*.

**2.2.4** En annan viktig faktor när det gäller urvalet är dimensionen *allmänspråk — fackspråk*. Gränsen mellan allmänspråkliga och fackspråkliga ordböcker är inte alltid helt knivskarp. Allmänt kan man säga att det är vanligare med renodlade fackordböcker än med renodlat allmänspråkliga ordböcker. Att avgöra vilka ord som inte är facktermer inom ett givet område är ju lättare än att avgöra vilka ord som är fackspråkliga och därför skall utmönstras ur en i sträng mening allmänspråklig ordbok.

Men på den här punkten — liksom fallet var med de synkroniska samtidsordböckerna — är det oftast användarkraven som får avgöra. Är t.ex. användarna av en "svensk-engelsk affärsordlista" nöjda med att bara få de svenska affärstermerna översatta till engelska? Eller vill de i stället ha en ordbok som i största allmänhet hjälper dem vid översättning av svenska affärstexter till engelska? I det senare fallet kan det finnas skäl att ta med allmänspråkliga ord och uttryck som ofta förekommer i affärstexter. På motsvarande sätt kan en i huvudsak allmänspråklig ordbok inte gärna underlåta att ta med de viktigaste av de oerhört många facktermer som dagligen används av gemene man.

**2.2.5** En annan dimension, som åtminstone delvis har med urval att göra, är ordbokens *attityd till språkbruket*. Den kan nöja sig med att redovisa existensen av ett visst språkbruk eller den kan påbjuda resp. motarbeta ett visst språkbruk. Vi får då motsatsparet *normativ* repektive *deskriptiv ordbok*.

De normativa ordböckerna kan vara *öppet normativa* eller *förtäckt normativa* (eller, med en vacker term som jag har hittat i sammanfattningen till ett av föredragen vid konferensen, *smugnativa*). De öppet normativa kan vara det på olika sätt, från att ge små diskreta kommentarer av normativ art i själva ordboksartiklarna<sup>3</sup>:

**Innovation** *f* Imponiervokabel für *Neuerung, Erneuerung, Entdeckung, Erfindung*

till att i förordet avge en programförklaring som denna:

"Das Wörterbuch der deutschen Gegenwartssprache ... wird den gesamten Wortschatz konsequent auf der Grundlage der marxistisch-leninistischen Weltanschauung darstellen. Das gilt für die Auswahl der Stichwörter, für die Bedeutungsangaben, die kommentierenden Bemerkungen und auch für die Auswahl der Beispiele."

Förtäckt normativa ordböcker är t.ex. sådana som har vinklat sitt urval av uppslagsord och exempel och som inte explicit anger principerna bakom urvalet. De kan på ytan se helt likadana ut som de deskriptiva.

Det finns också ordböcker som är delvis normativa, delvis deskriptiva. Det främsta exemplet i Sverige är väl *Svenska Akademiens ordlista*, som betraktas som norm för stavning och böjning av svenska ord. Urvalet av ord skall däremot inte uppfattas som normerande: Svenska Akademien har ju numera sprängt sexvallen i ordlistesammanhang och förtecknar även sådana ord som kan uppfattas som stötande, utan att för den skull vilja anbefalla dem till allmänt bruk.

Alla övriga allmänna ordböcker skulle man vilja sammanföra till en kategori med etiketten "normerande mot sin vilja". Ingen ordbok, den må vara hur deskriptivt menad som helst, kan ju helt undgå att uppfattas normativt.

### 3. Ordbokens förhållande till det språkliga tecknet

3.0 En viktig indelningsgrund för en klassifikation av ordböcker är dessas grundläggande *förhållande till det språkliga tecknet*, nämligen vilken komponent av tecknet — uttryck, innehåll eller funktion — de utgår ifrån:

- Ordböcker som baseras på lexikaliska enheters *egenskaper på uttryckssidan* brukar då kallas *semasiologiska* ordböcker.
- Ordböcker som baseras på lexikaliska enheters *egenskaper på innehållssidan* brukar kallas *onomasiologiska* ordböcker.
- Ordböcker som tar sikte på lexikaliska enheters *funktion*, dvs. deras sätt att bete sig i förhållande till andra enheter, kallas *syntagmatiska* ordböcker.

3.1 Bland de *semasiologiska* ordböckerna kan två huvudgrupper urskiljas.

3.1.1 Dels återfinns här den ordbokstyp som åtminstone lekmannen i första hand kommer att tänka på när han hör termen "ordbok": den *allmänna alfabetiska ordboken*.

3.1.2 Dels finner man här ett antal semasiologiska specialordböcker som redovisar *delaspekter av uttryckssidan*. Till denna grupp hör exempelvis *stavningsordböcker*, *avstavningsordböcker* och *uttalsordböcker*. Låter man ordens inbördes ordning grundas på ordsluten i stället för på ordens början, får man *baklängesordböcker* resp. *rimordböcker*. Hit hör också ordböcker över olika ord med identisk stavning, alltså *homografordböcker*. Till den här gruppen hör i själva verket en förfärande mängd olika ordbokstyper, från de mer professionella (t.ex. *frekvensordböcker* och *morfologiska ordböcker*) till sådana som är avsedda för korsordslösare och alfapetspelare (t.ex. ordböcker med orden ordnade efter antalet bokstäver, efter vissa bokstävers position i orden osv.).

3.2 Bland de *onomasiologiska* specialordböckerna kan man skilja mellan sådana som redovisar rent lexikaliska relationer och sådana där relationerna är mer associativa<sup>4</sup>.

3.2.1 Till dem som redovisar *lexikaliska relationer* kan man föra dels ordböcker över *synonymer* och över *antonymer*, dels ordböcker över *hyperonymer*, *hyponymer* och *kohyponymer*.

3.2.2 Till de ordböcker som redovisar mer *associativa relationer* skulle man kunna räkna t.ex. *tesaurer* och *bildordböcker*. Det handlar här om innehållsmässigt lösare förbindelser mellan orden. Det kan ibland räcka med att det föreligger ett visst mått av överensstämmelse mellan de situationer i vilka orden i fråga används.

En thesaurus av typ Roget's Thesaurus är naturligtvis en mer renodlat onomasiologisk ordbok än exempelvis en synonymordbok. I thesaurusen är orden redan från början systematiskt grupperade efter innehållsliga kriterier, och till det "associativa" hör även att thesaurusen är ordklassoberoende: det räcker med den begreppsmässiga släktskapen.

Bildordboken är också onomasiologisk i den meningen att den samlar ett antal begrepp inom ett visst ämnesområde, åskådliggjorda genom bilder, och anger deras benämningar på uttryckssidan.

**3.2.3** För att öka användbarheten brukar man ofta förse onomasiologiska ordböcker med en "semasiologisk ingång", så att säga, alltså ett alfabetiskt register, som ibland kan få avsevärda dimensioner. Den upplaga av Roget som jag har tillgång till består av 657 sidor thesaurus och 600 sidor register.

**3.3** *Syntagmatiska* specialordböcker är alltså sådana som tar sikte på de lexikaliska enheternas sätt att bete sig i förhållande till andra enheter. Det handlar här dels om valens, syntaktisk och semantisk, dels om urvalsrestriktioner på uttryckssidan. De ordbokstyper som kommer i fråga här är alltså:

- *konstruktionsordböcker*, som tar sikte på den syntaktiska valensen,
- *kollokationsordböcker*, som tar sikte både på den semantiska valensen och på urvalsrestriktionerna på uttryckssidan,
- *idiomordböcker*, som redovisar sådana syntagm som är så fasta både i fråga om syntaxen och ordvalet att det knappast går att ändra på någondera,
- *ordspråksordböcker* och *citatorordböcker*, där variationsmöjligheterna nästan är lika med noll.

#### 4. Antalet språk i ordboken

**4.1** En viktig indelningsgrund återstår, nämligen om ordboken är *enspråkig* eller *tvåspråkig*.

- I en enspråkiga ordboken förklaras ingångsspråkets lexikaliska enheter med hjälp av språket självt. Detta kan ske genom att man antingen anger *ersättare på uttryckssidan* (synonymer, parafrafer) eller ger mer renodlade *beskrivningar av innehållssidan* (definitioner).
- I den tvåspråkiga ordboken handlar det om att för varje lexikalisk enhet påkällspråket ange *motsvarigheten* (ekvivalenter) *på målspråkets uttryckssida*.

Av detta följer en annan viktig skillnad:

- I den enspråkiga ordboken är betydelsebeskrivningen inte nödvändigtvis insättbar på uppslagsordets plats.

- I den tvåspråkiga ordboken skall ekvivalenten vara insättbar på uppslagsordets plats.

Tvåspråkiga ordböcker behandlar som regel det normala samtida standardspråket. Det finns dock undantag. Det viktigaste är väl de *historisk-filologiska ordböckerna över döda språk*, t.ex. bibelhebreiska, antik grekiska och latin. Men egentligen borde man också kunna kalla en ordbok över en äldre variant av ett fortfarande existerande språk för tvåspråkig: man "översätter" ju till en yngre språkvariant. Samma sak skulle man nog taget kunna säga om ordböcker över särskilda varianter av samtidspråket, t.ex. dialektordböcker och slangordböcker: där "översätter" man ju på sätt och vis till standardspråket. Ibland kallas den här typen av ordbok också *internt tvåspråkig*.

En annan viktig skillnad mellan en- och tvåspråkiga ordböcker har att göra med användarnas förmodade kompetens:

- Den enspråkiga ordboken förutsätter att användaren har bristande kunskap om modersmålet.
- Den tvåspråkiga ordboken förutsätter att modersmålet är känt för användaren.

4.2 Den sistnämnda synpunkten är viktig för den grundläggande indelningen i tvåspråkiga ordböcker *till resp. från* modersmålet, alltså i *passiva resp. aktiva* ordböcker.

Distinktionen mellan passiva och aktiva ordböcker är oerhört viktig, eftersom den får helt avgörande betydelse för ordbokens utformning. Översättning kan ju ske i två riktningar, varvid användarens modersmål i det ena fallet är målspråk och i det andra fallet källspråk. För varje "översättningsriktning" är det bara vissa typer av språklig aktivitet som kommer i fråga, och därmed är det också bara vissa typer av information som behöver ges. Det är alltså av avgörande betydelse om modersmålet är källspråk eller målspråk<sup>5</sup>.

Den slutsats man då drar av allt detta är att det för ett givet par av språk behövs *fyra* ordböcker, för exempelvis språkparet svenska-franska alltså följande:

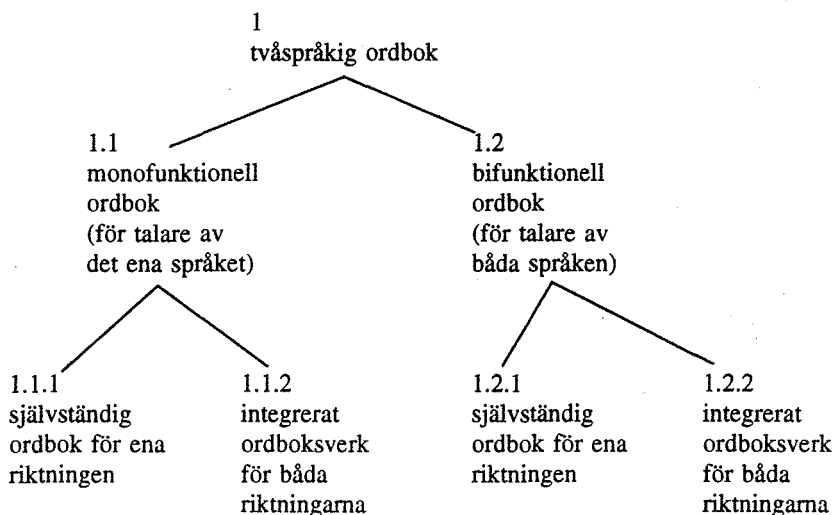
- för svenskspråkiga användare:
  - en "passiv" ordbok för reception av fransk text
  - en "aktiv" ordbok för produktion av fransk text
- för franskspråkiga användare:
  - en "passiv" ordbok för reception av svensk text
  - en "aktiv" ordbok för produktion av svensk text<sup>6</sup>

De finns också de som anser att det för varje språkpar behövs så mycket som *åtta* ordböcker<sup>7</sup>.

- De *passiva* ordböckerna indelas då i två kategorier, dels sådana som används för enbart *förståelse* av läst text ("Lesewörterbücher"), dels sådana som används för att *översätta* texten till modersmålet ("Herübersetzungswörterbücher").
- På motsvarande sätt skulle då de *aktiva* indelas i dels ordböcker för egen *textproduktion* på det främmande språket ("Schreibwörterbücher"), dels ordböcker för

översättning av en given text till det främmande språket ("Hinübersetzungswörterbücher").

4.3 Det här resonemanget bygger alltså på uppfattningen att en given tvåspråkig ordbok bör vara skraddarsydd för talare av enbart det ena språket; den skall vara *monofunktionell*. Motsatsen är då en *bifunktionell* ordbok, avsedd för talare av båda språken. Under båda dessa kan man finna två olika typer: dels den självständiga ordboken för bara en riktning, dels ett integrerat ordboksverk för båda riktningarna<sup>8</sup>. (Se Figur 3).



Figur 3.

Den självständiga ordboken i endast en riktning avsedd för talare av endast det ena språket (1.1.1) kan exemplifieras med Norstedts *Stora engelsk-svenska ordboken* eller med samma förlags *Stora svensk-engelska ordboken*. Ett integrerat ordboksverk för båda riktningarna men som är avsett enbart för talare av det ena språket (1.1.2) är van Dales nederländsk-tyska och tysk-nederländska ordböcker.

Den självständiga ordboken i endast en riktning avsedd för talare av båda språken (1.2.1) kan exemplifieras med Kjærulff Nielsen: *Engelsk-dansk ordbog* eller med Vinterberg/-Bodelsen: *Dansk-engelsk ordbog*. Typiska exponenter för typen integrerat ordboksverk för båda riktningarna avsett för talare av båda språken (1.2.2) är Collins' tvåspråkiga ordböcker, t.ex. den engelsk-tyska/tysk-engelska.

4.4 Det som nu har sagts om den tvåspråkiga ordboken leder tämligen automatiskt till slutsatsen, att en *tre- eller flerspråkig ordbok* med nämnvärda ambitioner i fråga om semantisk precision knappast är möjlig när det gäller allmänspråket: det är för mycket som inte överensstämmer mellan språken, både syntaktiskt och semantiskt, och ordbokstexten skulle bli oöverskådlig och användarovänlig. När det gäller fackspråk är förutsättningarna gynnsammare, i varje fall om man begränsar sig till översättning av enstaka termer och håller sig till fackområden som inte är kulturspecifika.

**4.5** Bortsett från de mer renodlade typerna — den enspråkiga ordboken på modersmålet och den tvåspråkiga ordboken till eller från modersmålet — finns det andra, där användarnas modersmål spelar en mer underordnad roll.

**4.5.1** En av dessa är den *enspråkiga ordboken avsedd för personer som inte har språket som modersmål* av typ *Advanced Learner's Dictionary, Longman Dictionary of Contemporary English* osv.

**4.5.2** En annan typ är den ordbok som har i princip samma målgrupp och utformning som den förra men där dessutom alla ord och uttryck översätts till användarnas modersmål. Det är alltså fråga om en *kombination av enspråkig och tvåspråkig ordbok*. Exempel på sådana ordböcker har vi på senare år börjat få i Sverige genom det s.k. LEXIN-projektet (LEXikon för INvandrare).

## 5. Informationskategorier i ordboken

**5.0** Jag har nu mer och mer avlägsnat mig från specialordböckerna och mer och mer koncentrerat mig på de allmänna alfabetiska ordböckerna, vilka alltså är de som lekmannen kommer att tänka på när han hör termen "ordbok". Det är nu dags att se efter vilka informationskategorier man kan påträffa i en sådan ordbok.

Även inom lexikografin finns traditionellt en helig treenighet, nämligen ordboksartikelns tre huvuddelar. Schemat är främst tillämpbart på enspråkiga ordböcker:

- Lemma
- Explikation
- Demonstration

**5.1** Under *lemmat* skulle då redovisas de kategorier som behövs för identifikationen av den lexikaliska enhet som skall beskrivas. Det handlar då i första hand om *stavning, uttal, böjning* och *ordklassstillhörighet*.

**5.2** Under *explikation* skulle falla dels *betydelsebeskrivning*, dels *paradigmatisk* information, dels *syntagmatisk* information.

**5.2.1** *Betydelsebeskrivningen* består av definitioner eller parafrafer. I tvåspråkiga ordböcker motsvaras betydelsebeskrivningen av ekvivalenter på målspråket. Betydelsebeskrivningen stöds ofta av uppgifter av icke-språklig natur, t.ex. encyklopedisk information — verbal eller visuell.

**5.2.2** Den *paradigmatiska informationen*, som främst är typisk för den enspråkiga ordboken, ges i form av begreppsbesläktade ord såsom *synonymer, antonymer, hyperonymer, hyponymer* och *kohyponymer*.

**5.2.3** Den *syntagmatiska informationen* utgörs av uppgifter om uppslagsordets sätt att fungera tillsammans med andra lexikaliska enheter. Här får vi alltså uppgifter om dels *konstruktioner*, dels *kollokationer*, dels *idiom*.

Men egentligen är det delvis fråga om flytande gränser här, eftersom uppgifterna om uppslagsordets paradigmatiska och syntagmatiska egenskaper åtminstone delvis kan sägas utgöra ett led i betydelsebeskrivningen.

**5.3** Avdelningen *demonstration* skulle då i huvudsak utgöras av *exemplen*. Man kan om man vill skilja mellan exempel som visar syntaktiska mönster, sådana som visar kollokationer och sådana som visar hela satser eller meningar. Exempelen i den sistnämnda kategorin är antingen autentiska (*belägg*) eller konstruerade av lexikografen (*redaktionsexempel*).

Också här är gränserna flytande, eftersom mycket av exempelvis den syntagmatiska informationen i praktiken ges i form av exempel.

**5.4** Dessutom anges genom olika typer av markeringar olika restriktioner i ordens och uttryckens bruklighet. Det kan gälla dimensioner som tid, rum, stilnivå, ämnesområde, frekvens och talarattityd men också motsatspar som bokstavligt—överfört och abstrakt—konkret:

åld.	hist.	dial.	Am.	poet.	vard.	sl.	vulg.	biol.	mus.
vanl.	ibbl.	iron.	neds.	eg.	bildl.			abstr.	konkr.

**5.5** I vissa enspråkiga ordböcker förekommer också en informationskategori som inte låter sig inordnas i någon av komponenterna i den lexikografiska treenigheten, nämligen *etymologisk information*. Längs tidsaxeln skär den genom de övriga kategorierna och upptar i sig element från flera av dem.

## 6. Ordbokens struktur

**6.0** De olika informationskategorierna måste nu inom ordboksartikelns ram arrangeras på ett ändamålsenligt sätt, och jag kommer nu över till den strukturella sidan. Det sätt på vilket ordboksartiklarnas delar förhåller sig till varandra brukar ju kallas *ordbokens mikrostruktur*.

På motsvarande sätt måste ordboksartiklarna i sin tur arrangeras i en inbördes ordning och med en viss layout som svarar mot ordbokens inriktning och användningsområde. Det sätt på vilket ordboksartiklarna förhåller sig till varandra brukar ju kallas *ordbokens makrostruktur*.

Mikrostruktur och makrostruktur är inga triviala frågor när det gäller att beskriva och klassificera ordböcker. De utgör viktiga element i betydelsebeskrivningen, dock på olika sätt och i olika hög grad i olika typer av ordböcker. Den enspråkiga ordboken är språkbeskrivande, dvs. den kartlägger språkets betydelsestruktur. Den tvåspråkiga ordboken har inte till syfte att beskriva ingångsspråkets betydelsestruktur utan skall i första hand möjliggöra översättning från ett språk till ett annat. En konsekvens blir att de strukturella frågorna spelar en viktig semantisk roll i den enspråkiga ordboken, medan de i vissa avseenden saknar betydelse i den tvåspråkiga.

**6.1** För att nu börja med *mikrostrukturen*, uppstår där dels problemet efter vilka kriterier indelningen i delbetydelser skall ske, dels problemet i vilken ordning de sålunda upprättade delbetydelserna skall presenteras.

### 6.1.1 Tänkbara kriterier för *indelning i delbetydelser* kan vara:

- *formella*: att en viss betydelse är begränsad till en viss form av ordet
- *syntagmatiska*: att en viss betydelse exempelvis förutsätter en viss syntaktisk eller semantisk valens
- *paradigmatiska*: att en viss betydelse exempelvis är kopplad till vissa synonymer eller antonymer
- *pragmatiska*: att en viss betydelse exempelvis är begränsad till ett visst ämnesområde eller en viss kommunikationssituation.

### 6.1.2 Också när det gäller den inbördes *ordningen mellan delbetydelserna* kan man grovt räkna med fyra principer:

- *historisk*: att man börjar med den äldsta betydelsen och slutar med den yngsta
- *logisk*: att man söker få fram en eller flera "huvudbetydelser" och visar hur de övriga sedan utvecklats därur
- *distributionsbaserad*, t.ex. efter syntaktisk distribution
- *frekvensbaserad*, vilket normalt innebär att man börjar med den vanligaste och slutar med den ovanligaste.

Var och en av dessa metoder har sina nackdelar, och det är egentligen inte någon av dem som ensam klarar av alla problem. Ofta måste lexikografen i praktiken tillämpa en kombination av flera av dem.

6.2 Det finns också beskrivningsproblem som ligger *mitt emellan mikro- och makrostruktur*, alltså då man måste välja om en viss företeelse skall avspeglas i mikrostrukturen eller i makrostrukturen.

6.2.1 Ett sådant är det klassiska problemet *polysemi kontra homonymi*, alltså om två lexikaliska enheter med samma form på uttryckssidan men olika betydelse på innehållssidan skall beskrivas som två olika betydelser hos "samma" ord eller som två "olika" ord.

- Bestämmer man sig för alternativet två betydelser hos samma ord är det alltså fråga om *polysemi*, och polysemi manifesterar sig i ordbokens *mikrostruktur*, vilket sker genom att de två lexikaliska enheterna presenteras som delbetydelser i en och samma artikel.
- Anser man i stället att det är fråga om två olika ord är det alltså fråga om *homonymi*, och homonymi manifesterar sig i ordbokens *makrostruktur*, vilket sker genom att de två lexikaliska enheterna presenteras i olika artiklar.



Olika ordböcker tillämpar här olika synsätt:

- Den traditionella metoden bygger på *språkhistoriska* kriterier. Lika ursprung innebär polysemi, olika ursprung betraktas som homonymi:

<sup>1</sup>bas -en -ar 'djup röst'; 'lägsta stämma'  
<sup>2</sup>bas -en -ar 'arbetsförman'  
<sup>3</sup>bas -en -er 1 'fundament'; 'underlag'; 'stödjepunkt' 2 'typ av kemisk förening'  
<sup>4</sup>bas (oböjl.) vard. 'stryk'  
<sup>1</sup>basa -de -t 'vara förman'  
<sup>2</sup>basa -de -t 1 vard. 'ge stryk' 2 'uppmjuka genom upphettning'  
basal -t  
basera -de -t  
basfiol  
basisk -t  
basist -en -er  
baslösning  
basning -en -ar 1 vard. 'kok stryk' 2 'uppmjukning genom upphettning'  
basvinkel

- Ett mer nutidsinriktat synsätt bygger på *semantiska* kriterier. Betydelser som kan härledas ur en gemensam huvudbetydelse innebär polysemi, medan helt skilda betydelser innebär homonymi:

<sup>1</sup>bas -en -ar 'djup röst'; 'lägsta stämma'  
<sup>2</sup>bas -en -ar 'arbetsförman'  
<sup>3</sup>bas -en -er 'fundament'; 'underlag'; 'stödjepunkt'  
<sup>4</sup>bas -en -er 'typ av kemisk förening'  
<sup>5</sup>bas (oböjl.) vard. 'stryk'  
<sup>1</sup>basa -de -t 'vara förman'  
<sup>2</sup>basa -de -t vard. 'ge stryk'  
<sup>3</sup>basa -de -t 'uppmjuka genom upphettning'  
basal -t  
basera -de -t  
basfiol  
basisk -t  
basist -en -er  
baslösning  
<sup>1</sup>basning -en -ar vard. 'kok stryk'  
<sup>2</sup>basning -en -ar 'uppmjukning genom upphettning'  
basvinkel

- En tredje metod innebär att man utgår från *formella* kriterier. Lexikaliska enheter som stavas lika men i övrigt har olika formella egenskaper betraktas som fall av homonymi och presenteras i olika artiklar, medan de som har samma formella egenskaper betraktas som fall av polysemi och presenteras i en och samma artikel, utan hänsyn till ursprung, betydelse eller morfologi:

<sup>1</sup>bas -en -ar 1 'djup röst'; 'lägsta stämma' 2 'arbetsförman'  
<sup>2</sup>bas -en -er 1 'fundament'; 'underlag'; 'stödjepunkt' 2 'typ av kemisk förening'  
<sup>3</sup>bas (oböjl.) vard. 'stryk'  
<sup>1</sup>basa -de -t 'vara förman'  
<sup>2</sup>basa -de -t 1 'vara förman' 2 vard. 'ge stryk' 3 'uppmjuka genom upphettning'  
 basal -t  
 basera -de -t  
 basfiol  
 basisk -t  
 basist -en -er  
 baslösning  
 basning -en -ar 1 vard. 'kok stryk' 2 'uppmjukning genom upphettning'  
 basvinkel

— En fjärde metod är den morfosemantiska, som alltså bygger på både *semantiska* och *morfologiska* kriterier. Den innebär att lexikaliska enheter med samma formella egenskaper men med helt olika betydelse kommer att anses utgöra exempel på homonymi, inte polysemi. Man så att säga "avlastar" mikrostrukturen och gör makrostrukturen mer komplicerad. Men till grund för indelningen har man också lagt ordens morfologiska egenskaper, vilket innebär att man under varje betydelse också tar upp olika avledningar och sammansättningar som är möjliga. En viss avledning kan alltså förekomma på flera ställen i ordboken:

<sup>1</sup>bas -en -ar 1 'djup röst'; 'lägsta stämma' - basist -en -er - basfiol  
<sup>2</sup>bas -en -ar 'arbetsförman' - <sup>1</sup>basa -de -t 'vara förman'  
<sup>3</sup>bas -en -er 'fundament'; 'underlag'; 'stödjepunkt' - basalt -t - basera -de -t  
 - basvinkel  
<sup>4</sup>bas -en -er 'typ av kemisk förening' - basisk -t - baslösning  
<sup>5</sup>bas (oböjl.) vard. 'stryk' - <sup>2</sup>basa -de -t 'ge stryk' - <sup>1</sup>basning -en -ar 1 vard. 'kok stryk'  
<sup>3</sup>basa -de -t 'uppmjuka genom upphettning' - <sup>2</sup>basning -en -ar 'uppmjukning genom upphettning'

6.2.2 Ett annat fall där man måste välja mellan mikro- och makrostruktur är då man står inför en *lexikaliserad enhet som består av mer än ett grafiskt ord*. Att i en engelsk ordbok hopskrivna sammansättningar som *airline* och *air-force* skall bli uppslagsord och alltså tillhöra makrostrukturen är väl helt klart, men hur är det med en helt analogt bildad men särskrivna sammansättning som *air officer*? Skall den redovisas under något av de enkla orden och alltså tillhöra mikrostrukturen? Att det franska *quoi que* skall tillhöra makrostrukturen är självklart, men hur är det med det analogt bildade och nästan synonyma men särskrivna *bien que*? På den här punkten tillämpar olika ordböcker olika principer.

6.3 När det slutligen gäller *makrostrukturen*, så kan man i de allmänna ordböckerna skilja mellan ordböcker med ett enda uppslagsord per stycke och ordböcker med flera uppslagsord per stycke.

**6.3.1** I de allmänna alfabetiska ordböckerna med ett uppslagsord per stycke blir det normalt fråga om strängt alfabetisk ordning, och varje uppslagsord kan sökas längst till vänster i spalten.

**6.3.2** I ordböcker med flera uppslagsord per stycke kan ordningen vara strängt alfabetisk men behöver inte vara det.

**6.3.2.1** Att sammanföra flera uppslagsord i ett stycke med bibehållen alfabetisk ordning kallas på engelska *niching*. Jag har förgäves sökt efter en bra svensk term: "gruppering i nischer" får duga tills vidare. De sammanförda uppslagsorden kan vara morfologiskt besläktade eller inte. Metoden medför inga sökproblem: serien av uppslagsord lämnar tillfälligt spaltens vänsterkant och rör sig i en slinga för att så småningom komma tillbaka till vänsterkanten igen, men den alfabetiska ordningen bryts aldrig. Förfarandet används oftast i utrymmesbesparande syfte. I små ordböcker kan det t.o.m förekomma att man sammanför samtliga uppslagsord i hela ordboken till ett och samma stycke<sup>9</sup>:

fliegen\* flyga **Fliegen** *n* flygning **Flieger** - *m* flygare **Fliegeralarm** - *e* *n* flyglarm **fliegen\*** fly, rymma **Fliese** -*n* *f* stenplatta **Fließband** -*er* *n* löpande band **Fließbandarbeiter** - *m* tempoarbetare **fließen\*** flyta, rinna, flöda; *fließend* *Deutsch sprechen* tala tyska flytande **flimmern** flimra **flink** kvick **Flinte** -*n* *f* bössa **Flirt** [-ö-] -*s* *m* flört **flirten** flörta **Flitterwochen** (*pl*) smekmånad **flocht** imperf. av *flechten* **Flocke** -*n* *f* ...

**6.3.2.2** Att sammanföra flera uppslagsord i ett stycke med avvikelser från den alfabetiska ordningen kallas på engelska *nesting*. Också här har jag förgäves sökt efter en bra svensk term: "gruppering i nästen" får duga tills vidare. De sammanförda uppslagsorden är morfologiskt besläktade, och skälet till att man sammanför dem är alltså pedagogiskt. Ordningen mellan huvuduppslagsorden i vänsterkanten är fortfarande strängt alfabetisk, medan de underordnade uppslagsorden ofta bryter den alfabetiska ordningen. Man kan på så sätt få en dubbel, tredubbel osv. makrostruktur i ordboken. Ett autentiskt svenskt exempel<sup>10</sup>:

nattvard ...

natur — natur//anlag ... -äng — natura

— natura//förmån ... -stipendium — naturalisera

naturalisering naturalism naturalist naturalistisk

naturell naturism naturist naturlig naturligtvis

nautisk ...

Här har vi faktiskt en fyrdubbel makrostruktur, med fyra alfabetiska serier att ta hänsyn till:

1 Huvuduppslagsorden (*nattvard*, *natur*, *nautisk*)

2 Sammansättningarna med *natur* som förled (*naturanlag* ... *naturäng*).

3 Avledningarna av *natur* (*natura* plus *naturalisera*, *naturalisering*, ..., *naturligtvis*).

4 Sammansättningar med avledningarna som förled (*naturaförmån* ... *naturastipendium*).

Det här är ändå ett ganska lindrigt fall från söksynpunkt, eftersom inget av orden vid strängt alfabetisk sortering skulle hamna före *nattvard* eller efter *nautisk*.

6.4 Till sist skall jag bara passa på att visa ett exempel på hur man kan spara utrymme genom att låta mikro- och makrostruktur gå in i varandra. Idén är inte ny. Exemplet är över 300 år gammalt<sup>11</sup>:

**Machiavelize**, the same as  
**Machiavelianize**, to play the  
**Machiavelian**, he that practiseth and studieth  
**Machiavelianism**, state policies, the doctrine of Nicholas  
**Machiavel**, a famous historian and recorder of Florence, whose politicks have  
poisoned almost all Europe.

## 7. Ordbokens förhållande till användarna

7.0 På ett par undantag när har jag hittills uppehållit mig vid den *fenomenologiska* aspekten av ordbokstypologin. Nu skall jag kort säga något om den *funktionella* sidan.

När det gäller ordböckers användning kan man om man vill skilja mellan två huvudgrupper: ordböcker för konsultation och ordböcker för inläring<sup>12</sup>.

7.1 Exempel på *ordböcker för inläring* skulle då vara skolordböcker, vokabelsamlingar, synonymordböcker, bildordböcker, ordböcker för andraspråksinläring.

7.2 Bland ordböcker för konsultation skulle man kunna skilja mellan dels ordböcker för allmänt bruk, dels ordböcker för arbete inom olika fackområden.

- ordböcker för textreception
- ordböcker för textproduktion
- översättningsordböcker.

7.2.2 Gruppen *ordböcker för arbete inom olika fackområden* kunde indelas i

- ordböcker för specialister på språket självt eller dess texter (lingvister eller filologer)
- ordböcker för specialister inom övriga fackområden.

## 8. Lexikografi och terminologi

Jag utgick i början från ordbokens förhållande till en besläktad produkt som har ett betydligt vidare område som objekt för beskrivningen, alltså *encyklopedin*. Avslutningsvis skall jag nu, beväpnad med den begreppsapparat vi har tillägnat oss, säga något om ordbokens förhållande till en besläktad produkt som har ett betydligt snävare område som objekt för beskrivningen, nämligen den *terminologiska ordlistan*. Jag menar då inte den vanliga fackordboken utan den typ av ordlistor som utarbetas enligt terminologilärans principer och utges av terminologi- och standardiseringsorganen.

Ser man på den vanligaste typen av kommersiell lexikografisk produkt, den allmänna alfabetiska ordboken, skulle man kunna säga att den är

- snarare semasiologisk än onomasiologisk
- snarare allmänspråklig än fackspråklig
- snarare för lekmän än för fackmän
- snarare deskriptiv än normativ
- snarare samtidsinriktad än historisk
- *antingen* enspråkig, och därvid både aktiv och passiv *eller* tvåspråkig, och därvid normalt antingen aktiv eller passiv.

Om man nu vill hårdra det hela en smula och kanske t.o.m. överdriva litet, skulle man kunna säga att det med den typiska terminologiska produkten ofta förhåller sig tvärtom. Den är

- *snarare onomasiologisk än semasiologisk*, genom att begrepp, begreppsrelationer och begreppssystem står i centrum för terminologen på ett helt annat sätt än för lexikografen. Den yttre manifestationen av detta finner man i ordlistornas struktur, som i sin mest renlärliga form är klart onomasiologisk, t.ex. genom att ha en makrostruktur som försöker reproducera det underliggande begreppssystemet. De enskilda termerna får man leta efter i ett alfabetiskt register.
- *alltid fackspråklig*, vilket ligger i sakens natur.
- *snarare för fackmän än för lekmän*, vilket också ligger i sakens natur.
- *snarare normativ än deskriptiv*, därför att terminologins uppgift är att skapa förutsättningar för en så entydig och konnotationsfri kommunikation som möjligt, medan den allmänspråkliga lexikografins uppgift är att beskriva ett faktiskt existerande språkbruk, inklusive konnotationer.
- *alltid samtidsinriktad*, vilket i praktiken följer av dess normerande inriktning: den blir därigenom inte bara *samtidsinriktad*, den blir också *framtidinriktad*.
- *både enspråkig och flerspråkig*, genom att de flesta terminologiska ordlistor har definitioner på minst ett språk och dessutom termekvivalenter på flera andra språk.
- *både aktiv och passiv*, genom att den begreppsliga överensstämmelsen mellan termer på olika språk i idealfallet är säkerställd och därför medger översättning i båda riktningarna.

Sammanfattningsvis och något tillspetsat:

- *Terminologins* huvuduppgift är att skapa ett rationellt medel för fackspråklig kommunikation.
- *Lexikografins* huvuduppgift är att beskriva ett redan förefintligt, ofta irrationellt medel för allmänspråklig kommunikation.

## 9. Avslutning

När jag läste sammandragen av de olika bidragen till konferensen, slog det mig att i stort sett alla de aspekter jag har berört kommer att täckas in under de närmaste dagarna — och därtill mycket annat. Det vittnar gott om mångsidigheten hos de lexikografiska aktiviteter som bedrivs på olika håll i Norden.

## Noter

1. Se vidare Lara 1989.
2. Se vidare Johnsson 1983:128.
3. Exemplet hämtat från Ripfel 1989:199.
4. Se vidare Reichmann 1990.
5. Se närmare Kromann et al. 1984:185 ff.
6. Även enspråkiga ordböcker kan naturligtvis beskrivas i dimensionen aktiv-passiv, beroende på hur mycket de inriktade på produktion respektive reception. Se t.ex. Svensén 1987:10-16.
7. Se t.ex. Hausmann 1977:144 ff.
8. Se Kromann 1987:11-13.
9. Exemplet hämtat från Prismas moderna tyska ficklexikon.
10. Exemplet hämtat från Prismas moderna tyska ficklexikon.
11. Coles: An English Dictionary ... (1676), citerat hos Osselton 1989:171.
12. Se vidare Kühn 1989.

## Litteratur

- Hausmann, Franz Josef. 1977. *Einführung in die Benutzung der neufranzösischen Wörterbücher*. Tübingen.
- . 1989. Wörterbuchtypologie. I: *Wörterbücher ...*, 1: 968-981.
- Illustrerad svensk ordbok*. 1977. Red. Bertil Molde. 3., rev. uppl. Stockholm.
- Jonsson, Hans. 1983. Om historisk lexikografi av typ SAOB. (*Nysvenska studier*, 63:127-158).
- Kromann, Hans-Peder. 1987. *Principles of bilingual lexicography*. (Manuskript ["vorläufige Endfassung"] till bidrag i *Wörterbücher ...*, 3.
- , Riiber, Theis & Rosbach, Poul. 1984. Überlegungen zu Grundfragen der zweisprachigen Lexikographie. I: H.E. Wiegand, red. *Studien zur neuhochdeutschen Lexikographie*, 5:160-238. Hildesheim.
- Kühn, Peter. 1989. Typologie der Wörterbücher nach Benutzungsmöglichkeiten. I: *Wörterbücher ...*, 1:111-127.

- Lara, Luis Fernando. 1989. Dictionnaire de langue, encyclopédie et dictionnaire encyclopédique: le sens de leur distinction. I: *Wörterbücher ...*, 1:280–287.
- Osselton, N.E. 1989. Alphabetisation in monolingual English dictionaries to Johnson. I: Gregory James, red. *Lexicographers and their works*, 165–173.
- Prismas moderna tyska ficklexikon*. 1974. Red. Eva Gomer & al. Stockholm.
- Reichmann, Oskar. 1990. Das onomasiologische Wörterbuch: Ein Überblick. I: *Wörterbücher ...*, 2:1057–1067.
- Ripfel, Martha. 1989. Die normative Wirkung deskriptiver Wörterbücher. I: *Wörterbücher ...*, 1:189–207.
- Svensén, Bo. 1987. *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. Stockholm. (Utkommer 1992 i engelsk version med titeln "Practical lexicography. Principles and methods of dictionary-making" [Oxford University Press].
- Wörterbücher, Dictionaries, Dictionnaires. Ein internationales Handbuch zur Lexikographie ...*, red. F.J. Hausmann, O. Reichmann, H.E. Wiegand & L. Zgusta. 1: 1989. 2: 1990. Berlin, New York.

## Ord på väg Några reflexioner kring nordisk lexikografi<sup>1</sup>

Det kan ha sina risker att ge ett bidrag som detta rubriken *Ord på väg*. Det låter oroväckande likt uttrycket *ord på vägen*. Några ord på vägen kan bara den komma med som har en överlägsen erfarenhet, och som anser sig kunna dela med sig av sin visdom. Det finns inget som helst underlag för att jag skulle ta på mig en sådan attityd. Snarare skall rubriken uppfattas i sin mera direkta betydelse, kanske någonting i stil med: det är en del på gång inom den språkvetenskapliga verksamhet som inriktar sig på ord, dvs. den lexikaliska forskningen och dess tillämpning inom praktisk lexikografi. Redan det faktum att det genomförs en nordisk lexikografisk konferens är ett tecken på att någonting håller på att hända. Dag Gundersen behandlar "leksikografi som undervisningsemne" i denna volym. En verksamhet som det regelrätt undervisas i, som det bedrivs forskning kring, som har en särskild, ganska entydig beteckning, och för vars skull man ordnar konferenser och ger ut tidskrifter, får också en egen identitet, verksamheten legitimeras, och vi kan få en akademisering av verksamheten.

### Lexikografins ställning

Det har sedan Dr. Johnson talats mycket om lexikografen som "a harmless drudge". Kanske håller lexikograferna åtminstone på att bli betydligt "less harmless drudges" nu. Så här skriver Zgusta i inledningen till sin bibliografi *Lexicography today* (1988): "Lexicography is one of the few areas in which linguistic activity has an immediate impact on many people, is very expensive (relative to the scale of measurement usual in the humanities) in terms both of time and money, and can be rather lucrative." Även om de flesta av oss inte märkt så mycket av den påstådda lönsamheten, så är vi nog överens med Zgusta för övrigt. Lexikografi är någonting som språkvetare kan syssla med utan att känna det som att man behöver ursäktas sig när andra människor frågar vad man gör. Så är det som bekant inte med alla aktiviteter inom det humanistiska området.

När en verksamhet tillåts dra resurser och uppfattas som praktiskt nyttig för folk i något avseende, så höjs också dess anseende normalt bland allmänheten. Nu har säkert lexikografen alltid varit en av de humanistiska verksamhetsgrenar som varit lättast att acceptera för så kallat vanligt folk. Det är väl snarast inom "språkfolkets" egna led som lexikografen ofta har viftats undan som en andra rangens sysselsättning. Vi vet ju att forskningsintensiv verksamhet lätt får högre status än exempelvis all slags undervisning, paradoxalt nog även vid våra högre utbildningsinstitutioner. Dessutom finns det en tendens

---

<sup>1</sup> Detta är en lätt reviderad skriftlig version av det anförande som jag hade i uppdrag att avsluta Oslo-konferensen med. I enlighet med uppdraget skulle jag presentera några personliga funderingar som konferensen hade givit upphov till. Denna förutsättning måste beaktas, eftersom den kan ha lett till en del formuleringar i det följande som kan verka förbryllande för de läsare som ej deltog i konferensen.



till att grundforskning och framför allt teoribygge värderas högre än tillämpad forskning och delvis också metodutveckling. Lexikografin riskerar att i det allmänna akademiska medvetandet reduceras till en verksamhet som på sin höjd tillämpar redan vunna forskningsresultat, som står i ett serviceförhållande till andra aktiviteter (som redan de åtnjuter mindre prestige, t.ex. undervisning), och som genom sina praktiskt baserade förutsättningar – t.ex. kravet att producera en ordbok före en viss tidsgräns – till och med de facto kan behöva avstå från mera sofistikerat teoretiserande.

När lexikografin nu tycks väcka ett ökande intresse, har det säkert att göra med det faktum att allt fler blir medvetna om behovet och nyttan av ordböcker. Lexikografiska hjälpmedel (framför allt ordböcker i traditionell mening) spelar en avgörande roll i tolknings- och översättningsarbete. Organiserad tolkning och översättning är i sin tur verksamheter som fått en alltmer central funktion i det moderna samhället, och den hela tiden tilltagande internationaliseringen skapar ett ständigt stegrat behov av dessa aktiviteter. Dessutom behöver översättnings- och tolkningsarbetet svara mot allt högre precisionskrav, till följd av den teknikalisering och i vissa avseenden komplicering som utmärker det omgivande samhället. Härigenom har tolkning och översättning – och därmed lexikografin – fått en erkänd samhällsfunktion.

Samtidigt har lexikografins ställning inom själva språkvetenskapens ram modifierats. Det finns antagligen många samverkande orsaker till att denna tendens uppstått. En kan sökas på metodsidan. I och med att lexikografiarbetet tog datorer i sin tjänst, kom man att associeras med en gren av språkvetenskapen, som i kraft av sin nära relation till den datamaskinella världen redan hade högt anseende, nämligen datalingvistik (i vid mening). När dessutom datalingvistik kom att koncentrera sina ansträngningar på att lösa lexikaliska problem inom ramen för sina speciella system, ändrades den datorbaserade lexikografins status radikalt. Datalingvistik har direkt efterfrågat lexikografisk information i välstrukturerad form, då sådan information behövs som en integrerad del av olika system för informationsökning, datorbaserad översättning etc.

En annan orsak till lexikografins ökande prestige är den lexikografiska teoribildning som med accelererande hastighet byggs upp inom verksamhetens egen ram. Det har länge hävdats från språkvetenskapligt håll att lexikografin kan utvecklas betydligt genom att man drar nytta av de resultat som vunnits inom den teoretiska lingvistik. Däremot har skiljelinjen mellan teori och praktiskt arbete alltid dragits så, att lexikografin hamnat uteslutande inom den praktiska tillämpningen. Även vid lexikografikonferenser brukar några föredragshållare nu för tiden poängtera skillnaden mellan bidrag som är av teoretisk natur och sådana som har ett mer praktiskt syfte, men då gäller det i allmänhet inte allmänlingvistisk teori gentemot lexikografisk praktik utan internt inom lexikografin.

Här får man kanske se upp, så att inte den gamla motsättningen mellan teoretiskt orienterade lingvister och (förment) praktiskt orienterade lexikografer inplanteras bland lexikograferna själva. Om några lexikografer utpekas som lexikografiteoretiker, medan andra, som kanske huvudsakligen arbetar med (den praktiska) framställningen av ordböcker, uppfattas som uteslutande praktiker (och därmed får lägre prestige i det allmänna omdömet), så har det tidigare problemet egentligen bara flyttats en aning. Målet måste vara att göra lexikografiskt arbete i gemen mer teorianknutet och hela tiden hålla det teoretiska utvecklingsarbetet i god kontakt med den praktiska erfarenheten. I själva verket borde lexikografi vara ett område där det *aldrig* skulle behöva råda någon motsättning mellan teori och praktik. Syftet är definitionsmässigt praktiskt – det är ordböcker eller liknande *produkter* som skall framställas – och målet nås lättare och med mer tillfredsställande resultat om praktiken vilar på en genomtänkt teori; samtidigt kan teorin oavbrutet hämta näring ur den

praktiskt betingade erfarenheten. På så vis skulle lexikografin närmast kunna tjäna som en idealisk förebild för många andra humanistiska verksamheter som har dåligt utvecklad teoribildning. Den ömsesidiga misstänksamheten mellan påstådda teoretiker och praktiker är vanligen helt onödig, men dess uppkomst är också lätt att förstå. Teoretikerna framhåller alltför ofta behovet av teori utan att vara särskilt väl förtrogna med det praktiska arbetet, och praktikerna flyr då gärna in i en teorifientlig attityd, och låter "teorier" bli liktydigt med "spekulationer". Vad som behövs för att närma de två grupperna till varandra är bland annat en avdramatisering av teoribegreppet, och det sker lättast om båda sidor inser nyttan av en integrerad syn. Jag tror det är detta som håller på att hända inom lexikografin.

### Lexikografi som språkvetenskap

Ytterligare en orsak till lexikografins ökade status är det faktum att dess ställning som språkvetenskaplig verksamhet håller på att ändras radikalt. Den språkteoretiska ståndpunkten har ju egentligen sedan Bloomfield gått ut på att lexikonet är ett "appendix of the grammar, a list of basic irregularities" (Bloomfield 1933, s. 274). Med Chomskys övertagande av detta lexikonbegrepp och genom den generativa grammatikens dominans inom teoribildningen blev lexikaliska frågor naturligen mycket dåligt uppmärksammade. När grammatikmodellen avbildades, återgavs den länge som en gigantisk syntaxkomponent med små sekundära komponenter för fonologisk representation respektive semantisk uttolkning kopplade till sig. Om lexikonet över huvud taget antydades, så var det som en egen liten ruta vid sidan av, förbunden med modellen i övrigt genom en streckad linje, på sin höjd.

Det kanske är överord att påstå att vi nu istället har en stor lexikonkomponent mitt i figuren, med syntaxen, morfologin (som åter kommit till heders), fonologin och semantiken ordnade som satellitkomponenter runtom, opererande på den centrala lexikonkomponenten. Men en tendens åt det hållet kan märkas; i alla händelser uppfattas de lexikaliska frågorna sedan en tid som betydligt mer centrala än tidigare. Det börjar också påpekas allt oftare att förhållandet mellan grammatiken och lexikonet i språkbeskrivningen mest är en perspektivfråga. Man kan välja att börja med att beskriva de grammatiska reglerna och se lexikonet som den idiosynkratiska information som måste specificeras särskilt för att göra beskrivningen fullständig.

Man kan emellertid också utforma lexikonet som en innehållsrik och välstrukturerad informationsmängd och extrahera de flesta grammatiska strukturer ur denna information. Det är ofrånkomligt att vissa grammatiska fenomen lättast beskrivs i allmänna grammatiska termer medan andra bäst behandlas i förhållande till speciella (klasser av) lexikaliska enheter och därför kan anses vara en del av den lexikaliska informationen. Det finns inget som säger att det finns en naturlig uppdelning mellan grammatik och lexikon, som det gäller att finna fram till, utan det är uteslutande en fråga om hur vi vant oss att se på dessa frågor. Det har t.o.m. hävdats att det är de traditionella kommersiella ordböckerna och grammatikerna som styr vår uppfattning på den här punkten.

Under senare tid har således språkvetenskapen alltmer kommit att ägna sig åt lexikaliska frågor. Därvid har de traditionella ordböckerna fått en ny roll som inspirationskälla till den teoretiska lingvistik. I samband härmed har också språkteoretiker modifierat sin tidigare onyanserat förenklade syn på lexikografiskt arbete. Man har helt enkelt upptäckt att informationen i ordböcker ofta bygger på en sofistikerad analys och – inte minst – avancerad materialkännedom. På den senare punkten kan man för närvarande märka en tydlig tendens till omsvängning i språkvetenskapen på det internationella planet. Långt generaliserade

påståenden på grundval av ett mycket begränsat material är inte lika gångbara längre. Man vänder i ökande omfattning tillbaka till undersökningar av stora material, och det anses åter vara en angelägen uppgift för datalingvister att utveckla stora – nu *mycket* stora – textkorporusar. Respekten för lexikografiskt arbete tycks tillta i takt med denna allmänna attitydförändring.

Mest direkt utnyttjade blir ordböcker i den verksamhet som ibland kallas "computational lexicology". Denna beteckning avser den gren av datalingvistik som framför allt sedan Amsler (1980) systematiskt lagt ordböcker till grund för språkvetenskapligt analysarbete i syfte att ur ordböckernas informationsmängder extrahera återkommande lexikaliska strukturer. Även om ordböcksarbetet ofta, kanske normalt, bedrivs av språkvetare som haft begränsat intresse för teoretisk lingvistik, så har ordböckernas information genomarbetats efter en medveten plan. Efterhand har också traditioner utvecklats, så att en ordbok som ansetts ha varit framgångsrik (t.ex. för att den just varit baserad på ett medvetet, teoretiskt genomtänkt grundarbete) kommit att bli mönsterbildande för nya ordböcker.

Med den här enkla rundmålningen har jag inte avsett att ge någon pretentiös beskrivning av "the state of the art". I så fall skulle betydligt mer behöva sägas. Men jag har velat understryka att dagens lexikografi har relevans på en hel rad olika sätt vid sidan av de traditionella. Lexikografi (i vid mening) har *inomvetenskapligt* kommit att bli en betydligt mera central språkvetenskaplig verksamhet; den är en viktig faktor både inom den teoretiska lingvistik och inom den mer praktiskt inriktade språkvetenskapen, och dessutom som en komponent i en rad språkvetenskapliga tillämpningar av stor kommersiell och samhällelig betydelse. Allt detta har bidragit till att höja lexikografins anseende, vilket är glädjande på ett allmänt plan för oss som i växlande omfattning sysslar med lexikografi. Men det medför också ett ansvar – det är vår skyldighet att vidmakthålla området betydelse och utveckla verksamheten ytterligare.

Även om lexikografi är en gammal väletablerad verksamhet, är det alltså en ung vetenskap, kanske egentligen snarast under framväxande. Det märks också på en del av bidragen i denna volym. En del författare är noga med att understryka behovet av framför allt teoriutveckling och teoretiska ställningstaganden, andra vill gärna framhålla att de mera ser sig själva som praktiker. När allt kommer omkring är de flesta bidragen väl förankrade åt båda hållen, såväl de som annonseras som rena fallstudier som de som tar upp mer principiella frågor, låt vara att tyngdpunkten kan falla lite olika. Det är nog så att modern lexikografi är ganska omöjlig att bedriva framgångsrikt utan teoretisk bas. Vi behöver nog inte vara så ängsligt vaksamma och ursäkt oss hela tiden.

Unga vetenskaper brukar besväras av olika avgränsningsproblem å ena sidan och av terminologiproblem å den andra. Den lexikografiska terminologin har till stor del uppstått spontant, och någon fullständigt enhetlig bild finns inte (se vidare Bo Svenséns bidrag i denna volym). Redan ordet *lexikon* används i åtminstone tre olika betydelser:

- (1) uppslagsverk;
- (2) ordbok, särskilt tvåspråkig ordbok;
- (3) människans ordförråd, särskilt i teknisk mening, den språkliga beskrivningen av människans ordförråd, eller den lexikaliska komponenten av en språkvetenskaplig beskrivningsmodell.

Egentligen borde ordet *lexikografi* vara mångtydigt på ett parallellt sätt, beroende på vilket slags *lexikon* enligt ovan man skriver (eller beskriver). Jag är inte alldeles säker på att *lexikografi* och *lexikon* står i ett fullständigt symmetriskt förhållande till varandra.

Vid sidan av termen *lexikografi* förekommer *lexikologi*, som också används svårartat mångtydigt. Olika användare kan med *lexikologi* mena åtminstone följande:

- (1) läran om lexikaliska frågor på det mest övergripande planet (dock oftast utelämnande uppslagsverk som inte i första hand lämnar språkliga uppgifter);
- (2) unionen av "praktisk" lexikografi och det teoretiska lexikografiska grundarbetet;
- (3) enbart lexikografiteori, dvs. teoretiskt arbete gällande ordboksframställning;
- (4) studiet av den lexikaliska komponenten inom en språkvetenskaplig teori;
- (5) lexikalisk semantik specifikt.

Om graden av terminologisk stabilitet röjer en vetenskapsgrens mognad, så är det uppenbart att det lexikaliska arbetet har en lång väg kvar att vandra. När det gäller termen *lexikologi* har det lanserats en begreppsutredning efter konferensen i Oslo. Martin Gellerstam har med hänvisning till en idé presenterad av Sven-Göran Malmgren (båda vid Göteborgs universitet) föreslagit att *lexikologi* i första hand skall ses som samlings termen för "allmän lexikalisk teori" och "lexikografi". Under "allmän lexikalisk teori" sorterar sådant som "lexikalisk morfologi", "lexikalisk syntax", "lexikalisk semantik", "studiet av ordförrådets struktur" i olika avseenden etc. Under "lexikografi" finner man "lexikografiteori" och "lexikografisk tillämpning".

Med hjälp av sådana här begreppsutredningar kan vi göra klart för oss vad vi menar, och avgränsningsfrågorna förenklas. Avgränsningen i sin tur får naturligtvis inte bli en fråga om revirtänkande och protektionism. Det är emellertid viktigt – mer positivt uttryckt – att vi kan definiera vårt ansvarsområde och utveckla vår professionalism. Man måste kunna kräva av oss att vi har en seriös inställning, och vi själva skall bland annat ha rätt till yrkesstolthet. Men avgränsningar förblir naturligtvis ett ständigt problem inom vetenskaplig verksamhet, på olika nivåer.

Nu spelar sådana här praktiskt betingade uppdelningar sällan någon roll för den språkbrukande människan, som är en total kunskapsvarelse som associerar fritt mellan olika typer av kunskap och erfarenhet. Att språkmodellen brukar fördes med två axlar – en mera utpräglad inomspråklig axel och en som i olika avseenden vetter mot världen – bekymrar normalt inte språkbrukarna i gemen. Man kan fråga sig hur mycket av de olika avgränsningar man brukar göra som är uteslutande inomvetenskapligt motiverat, eller helt enkelt traditionsbundet.

Över huvud taget är traditionen en av de verkligt starka krafterna inom lexikografin. Själva den information som ges i ordböcker dikteras ofta av traditionen – både i enskilda detaljer och i fråga om urval och övergripande organisationsprinciper. Denna sanning har framkommit vid en rad lexikografihistoriska undersökningar. Det är oftast ett fåtal förebilder som dominerar den lexikografiska utvecklingen inom ett språkområde. Mina egna lexikografihistoriska erfarenheter härrör framför allt från projektet "Lexikografiska traditioner i Sverige" (för en översikt se t.ex. Hannesdóttir & Ralph 1988), men de intryck vi fått stöds av andra liknande undersökningar. Det är traditionen snarare än systematiskt teoretiskt arbete som drivit lexikografin framåt.

### Den tekniska utvecklingen

Som tidigare antytts har tillkomsten av olika datorbaserade informationslagrings- och informationssökningssystem dragit in språkvetenskapen i ett helt nytt sammanhang. Därvid

har den lexikaliska sidan av språket hamnat i fokus för intresset. Lexikografi har helt enkelt blivit en erkänt central verksamhet, och den information som produceras är eftertraktad, kommersiellt användbar och samhällsrelaterad viktig. Men teknikens utveckling har en avgörande betydelse för lexikografin även på ett annat sätt.

Det kan vara angeläget att lösa avgränsningsfrågorna av en rad olika skäl, som vi såg nyss. Det är emellertid viktigt att inte de teoretiska frågorna överdramatiseras i det sammanhanget. Om det är så att många avgränsningar görs mest av traditionsskäl eller av praktiska skäl (exempelvis för att vi måste ta itu med en sak i sänder), så gäller det att inte upphöja dessa skäl till några överordnade, teoretiskt nödvändiga skäl.

Frageställningar som rör den tekniska utvecklingen har ingen framträdande plats i denna volym, vilket i och för sig är naturligt; vi är lexikografer, inte tekniker. Jag är verkligen inte rätt person att tala om teknik, men låt mig komma med några reflexioner. Många av de avgränsningsproblem som lexikografer (och andra språkvetare) har varit sysselsatta med kan möjligen mycket snart visa sig vara rena skenproblem. Så länge man måste koncentrera sig på att framställa en produkt i taget, av praktiska skäl, måste man ta ställning till innehållet i detalj. Den nya tekniken öppnar helt andra perspektiv än vi varit vana vid. Jag nöjer mig här med att hänvisa till exempelvis Rolf Gavares artikel i *Språkvard* 1/1991 om "Lexikaliskt, datatekniskt stöd vid översättning och författande" och Gavares, Heggstads och Norling-Christensens bidrag till rapporten från konferensen "LEXIKONORD" i Göteborg 1990.

Det är helt klart att man nu kan upprätta verkliga superdatabaser, som i någon mening skulle lösa de problem vi talat om. Vi behöver inte nödvändigtvis göra en strikt boskillnad mellan ordboksinformation och encyklopedisk information, om båda typerna kan lagras tillsammans. På samma sätt kan allmänspråk och fackspråk rymmas ihop, specialordböcker och universalordböcker etc. En människa gör ingen skillnad på den information vi brukar inkludera i enspråkiga ordböcker och sådan som snarare återfinns i två- eller flerspråkiga. På liknande sätt är aktiva eller passiva ordböcker snarast en fråga om olika perspektiv på samma information. Över huvud taget blir informationsurvalet mindre betydelsefullt ju mer man kan hålla åtkomligt som en sammanhållen informationsmängd. Man kan också fundera över om egennamn och termer är så väsensskilda enheter jämfört med annat lexikaliskt material, eller om det bara har blivit tradition att vi behandlar namn och termer för sig. Strängt taget är det mera naturligt att betrakta namn och termer som något extrema lexikaliska enheter.

När det gäller egennamn kan man göra en reflexion i kvantitativa termer. Ett av de mera genomarbetade språkliga material som jag känner till är den tidningspråkskorpus som ligger till grund för *Nusvensk frekvensordbok* 1-4 (1970-80). I den undersökningen omfattade korpusen drygt 1 miljon löpande ord. Ungefär 45.000 av dessa var egennamn (på grafordsnivån; enheter som blev egennamn först på frasnivån negligeras här). Ca 4,5 % av den löpande texten upptogs således av egennamn. Genom lemmatisering reducerades miljonordskorpusen till ungefär 71.000 olika lemman (uppslagsformer). Runt 11.000 av dessa hade status som egennamn. Den andel av texten som egennamnen utgjorde var således 4,5 %, men egennamnens sammantagna lexikaliska frekvens var så hög som ca 15 %. Från lexikaliska utgångspunkter finns det egentligen ingen anledning att utesluta egennamn ur ordböcker, såsom brukar ske. Tvärtom utgör de en proportionellt sett mycket viktig ordkategori. Däremot förblir det naturligtvis ett teoretiskt och praktiskt problem att välja ut de egennamn som skulle kunna ingå i en ordbok eller en lexikalisk databas (jfr Botolv Hellelands bidrag i denna volym).

På senare år har det uppstått ett praktiskt behov av genomarbetade terminologiska system (såsom oljeterminologin i Norge). Därigenom har en intensiv terminologisk/terminografisk verksamhet kommit igång. Från terminologihåll understryks ibland distinktionen

mellan terminologi/terminografi å ena sidan och lexikologi/lexikografi å den andra. Från rent lexikaliska eller språkvetenskapliga utgångspunkter är det inte lätt att motivera en sådan hållning. Utpräglade termer och utpräglade allmänspråkliga ord kan sägas gruppera sig kring ändpunkterna på en skala, där ordförrådet i övrigt fördelar sig över det mellanliggande utrymmet. Även allmänspråkets ord kan variera i precision och exakthet i betydelsen, och det är väl just den *reglerade* inordningen i ett terminologiskt system som skiljer termer från andra ord. Men detta i sin tur är inte märkvärdigare än att ordförrådet även på andra sätt är stratifierat i olika avseenden. Det är givet att terminologer har en speciell uppgift och av olika skäl kan behöva hävda terminologins autonomi åt olika håll, men termer är naturligtvis ett slags ord och därmed av lexikografiskt intresse.

Liksom för namnens del kan man säga att åtminstone det praktiska problemet med avgränsning upphör om man kan ta med allt i samma databas. När tekniken numera gör det principiellt möjligt, kan jag inte se annat än att lexikografen (i vidaste mening) därigenom tar ett veritabelt kvalitativt språng. Jag avstår emellertid på goda grunder från att gå in i några mer detaljerade tekniska utläggningar.

### Lexikografi och teoriutveckling

Om det nu är så att man kan lagra i stort sett allting i samma superdatabas – traditionellt lexikografisk information vid sidan av encyklopedisk, enspråkig och två- eller flerspråkig lexikografisk information, ljud- och bildillustrationer (även rörlig bild) som komplement till textbaserad information, allmänspråklig och specialspråklig information med mera, med mera – är det då så att den lexikografiska teorin blir överflödigt?

Naturligtvis inte. Tvärtom, frestas man nästan att säga. Fortfarande kvarstår mängder av teoretiska frågor som gäller organiseringen av informationen. En grundläggande, språkvetenskapligt intressant fråga är exempelvis i vilken utsträckning man skall sträva efter att återspegla människans lexikon i databasen. I så fall: Vems lexikon? Är det den genomsnittliga språkbrukaren som skall approximeras, och, om så, hur är en sådan skapt? Eller är det en kollektiv kompetens som skall modelleras, en kumulativ kunskapsmassa som är allas och ingens? Frågorna är många och intrikata.

Åtskilliga av de problem som tas upp i denna volym har med "den lexikaliska enheten" att göra. Det är en avgörande teoretisk fråga för både lexikografen och språkteorin i allmänhet. Den lexikaliska enhetens form och representation ger en ordbok ett av dess mest karakteristiska drag och påverkar hela organisationen av ordboken i övrigt.

Det morfologiska systemets relation till den lexikaliska enhetens natur är en besläktad fråga. Har (grund)morfem någon autonom lexikalisk funktion, eller är de bara en abstraktion? (Att avledningsmorfem anses ha karaktären av självständiga lexikaliska enheter tycks framgå av deras behandling i flera moderna ordböcker.) Hur hanterar man bäst lexikaliska uppgifter som hänför sig till enskilda böjningsformer snarare än till uppslagsordet som helhet? På samma sätt kan man fråga sig vilka syntaktiska uppgifter som skall anses lexikografiskt relevanta. Det är exempelvis vanligt att ange vissa former av verbvalens. I den korpusundersökning som återopades ovan hade ett dussintal ord, som alla klassificerats som tillhörande samma ordklass (adverb), och som hade jämförbar frekvens, helt olika egenskaper med hänsyn till benägenheten att ingå i fraser. Det såg ut som i tabellen (efter uppgifter i NFO 3 (1975)).

	<i>frekvens</i>	<i>varav i fraser</i>
verkligen	297	2
däremot	296	8
dessutom	346	13
knappast	307	39
ytterligare	291	99
snart	288	107
alltför	288	144
bort	298	179
ner	286	183
helst	303	215
tilbaka	333	239

Det finns naturligtvis olika slags språkstatistiska iakttagelser att göra här, och jag vill inte komma med några kraftfulla påståenden, men man kan i all enkelhet fråga sig om den här typen av information hör hemma i en lexikografisk databas. Det är ju trots allt intressant att veta om ett ord att dess förekomst i en text nästan alltid betingas av vissa andra ords förekomst, medan andra grammatiskt jämförbara ord uppträder mera självständigt. Hur skall över huvud taget kopplingen mellan den lexikaliska informationen och textperspektivet åstadkommas?

Här finns betydligt fler frågor än svar, och man kan därför bara upprepa maningen att fortsätta det på stabil lexikografisk grund baserade teoriarbetet. En teoretisk fråga som är överlagrad det mesta vi gör, och som dyker upp på många ställen och i många skepnader i de olika bidragen ovan, och som behöver en egen livlig diskussion, är *lexikaliseringproblemet*. Vad skall man mena med lexikalisering, och när är en lexikalisk enhet lexikaliserad? Här kan frågan bara ställas – det dröjer antagligen innan några acceptabla svar kan presteras.

Frågor som dessa aktualiserar i sin tur en hel rad traditionella språkvetenskapliga frågeställningar, som aldrig fått något definitivt, tillfredsställande svar. Egentligen återkommer vi ständigt till hela frågan om hur det produktiva lexikaliska regelsystemet är utformat: Vi tror kanske att vi vet hur ord kan bildas i vid mening, men hur skapas egentligen metaforer mera exakt? Vad spelar fraser, idiom, kollokationer för roll i ett lexikaliskt perspektiv? För att än en gång återvända till tidningsspråkskorpusen, så visade det sig att ett stort antal satsinledande block som *inte* var välformade konstituenten återkom med varierande, men i några fall hög, frekvens (*NFO 3 (1975)*, t.ex. listorna under "1. Ordkombinationsnivå"). Man kan fråga sig huruvida det rör sig om medvetna eller omedvetna, inlärda textinledningsstrategier; kan dessa i så fall betraktas som en del av vår lexikaliska kunskap?

Till problemkomplexet med lexikalisering och den lexikaliska enhetens natur hör också sådant som ords (lexikaliska enheters) grad av genomskinlighet och förutsägbarhet (i morfologiska termer), där man kan urskilja olika typologiska profiler för olika språk. Det rör sig exempelvis om i vilken utsträckning man utnyttjar helt arbiträra simplexord, morfologiskt förutsägbara bildningar som sammansättningar och avledningar, flerordskombinationer, lånord, (andra) morfem av främmande ursprung etc. Dessa frågor är givetvis relaterade till varje ordboks attityd till normering m.m.

Det säger sig självt att dessa påståenden inte innebär några häpnadsväckande upptäckter. Vi är alla medvetna om att de nämnda problemen mera har karaktären av att vara ständiga följeslagare, nästan i klass med de (andra) existentiella frågorna. Men tillfället är nu gynnsammare än kanske någonsin tidigare att komma en bit på väg med dessa frågor. Även

så till vida kan det vara motiverat att tala om "ord på väg"! Antagligen kan man komma bäst tillrätta med dem om man kombinerar ett allmänlexikalskt perspektiv med ett rent lexikografiskt. Det verkar många vara inställda på. De språkvetiker som nu närmar sig de lexikaliska frågorna intar inte samma nedvärderande hållning gentemot stora lexikografiska datamängder som deras föregångare ofta gjort. Samtidigt är lexikograferna i normalfallet välorienterade i allmän språkvetenskap nu för tiden. Till detta kommer som sagt en intressant teknisk utveckling som öppnar möjligheter som inte funnits förut.

Det kanske allra mest lovande är ändå att myndigheter och opinionsbildare visar ökad förståelse för språkfrågornas betydelse. Det är många som insett exempelvis vilka magnifika översättningsuppgifter som väntar med det tilltagande internationella samarbete som signalerats. Översättning ställer stora krav inte minst på de lexikografiska hjälpmidlen.

Jag passar på tillfället att nämna att humanistiska fakulteten vid Göteborgs universitet beslutat om en särskild satsning på ett lexikologiskt forskningsprogram. Med lexikologisk forskning har man då framför allt menat sådan forskning som gäller den lexikaliska komponenten av språkmodellen jämte lexikografiteori. Däremot hamnar den rent praktiska lexikografin, i betydelsen lexikografisk produktion, utanför. Liknande tendenser märks också i de andra nordiska länderna. Under 1991 har det också inrättats en tjänst i lexikalisk forskning vid det svenska Humanistisk-samhällsvetenskapliga forskningsrådet. Tjänsten lockade tretton högt kvalificerade sökande, vilket säger något om ämnesrådets aktualitet och vitalitet. Jag nämner bara dessa omständigheter, som jag råkar känna väl till, som exempel på att något håller på att hända på det lexikaliska området.

Även Nordiska samarbetsnämnden för humanistisk forskning (NOS-H) har beviljat medel till ett större nordiskt samarbetsprojekt inom lingvistik och datalingvistik med lexikonet som huvudsakligt undersökningsobjekt. Enligt planerna skall detta projekt för det första beskriva teoretiska regelbundenheter i de nordiska språkens lexikon; för det andra skall man utforma datamaskinella lexikon för de (i projektet inblandade) nordiska språken, så att lexikonerna kan utnyttjas i maskinell analys av satser och texter på dessa språk; för det tredje vill man samordna olika teoretiska lexikonprojekt av liknande slag. Till de nämnda satsningarna skulle ytterligare flera kunna läggas, men dels skulle det leda för långt, dels hänvisar en del andra bidrag i volymen till sådana projekt. Det står i alla händelser klart att det är åtskilliga lexikaliska projekt på gång i Norden, som garanterar att det är "ord på väg".

### Nordiskt perspektiv

Vad är nu nordiskt i allt detta? Som framgått är en del satsningar gemensamma nordiska samarbetsprojekt, finansierade med nordiska medel. Tillåt mig också peka på Nordiska ministerrådets förslag till handlingsprogram för språkligt samarbete i Norden 1990–1995, som bl.a. går ut på att en arbetsgrupp tillsätts med följande uppdrag (Nordiska rådet, 38:e sessionen 1990, Förtryck B 103/k, s. 35):

1. Att på grundval av den översikt över existerande ordböcker som presenterats i språksekretariatet 1988 och ny inventering av relevanta pågående ordboksprojekt utreda och värdera de återstående behoven av gedigna ordböcker mellan allmänspråken i Norden. [...]
2. Att ge förslag på hur kommande nordiskt ordboksarbete bäst skulle kunna organiseras och finansieras. Man bör alltså beakta vilka vinster som kan göras



i tid och arbete genom samordning, såsom upprättande av gemensamma databaser etc. Den primära inriktningen skulle vara ordböcker mellan språken i Norden, men på längre sikt kunde även gemensamma ordboksprojekt mellan språk i Norden och språk från mera fjärran länder bli aktuella.

Denna text förefaller mig vara både insiktsfull och bjudande. Det ankommer naturligtvis på oss som sysslar med lexikografi att lämna förslag till konkreta åtgärder. Jag påminner mig i sammanhanget en idé, som Baldur Jónsson framförde i *Språk i Norden* (1982, s. 13–14):

Med tanke på hur beroende vi är av goda ordböcker har det länge förvånat mig att ingen tagit upp tanken på en nordisk ordboksredaktion som en permanent institution med stöd från alla de nordiska länderna. Arbetet är behövligt; det är för stort för enskilda personer, men om man kan gripa sig an det med djärvhet och vidsynthet öppnar det möjligheter för rationella och harmoniserade arbetsmetoder.

Antagligen är det meningslöst att i dagens läge önska sig en gemensam nordisk ordboks-institution. Däremot är det i högsta grad angeläget att "det nordiska ordboksarbetet" samordnas. Språken är inte mer olika varandra än att det är en god hjälp att ha tillgång till lexikografisk information på ett av språken när man sysselsätter sig med ett av de andra. Det gäller i olika omfattning såväl de centralskandinaviska och de andra nordgermanska språken som de övriga, icke-germanska språken i Norden – kulturlikheterna och de allmänna överensstämmelserna i fråga om kommunikationsvanor är trots allt så långtgående att det färgar av sig även på den lexikaliska strukturen i stort.

Det nordiska samarbetet kan således ta sig olika uttryck. I vissa fall behövs det ordböcker (i någon form) som behandlar par av språk i Norden. I andra fall kan en existerande ordbok som involverar ett språk i Norden och något annat språk tjäna som förebild för en ny ordbok som översätter mellan samma icke-nordiska språk och ett annat språk i Norden än i det förstnämnda fallet. För vissa språk i Norden behöver helt enkelt enspråkiga databaser upprättas, och det är då naturligt att vända sig till andra nordiska språkområden för att söka förebilder. Under alla omständigheter kan teori- och metod-utveckling, utbyte av erfarenheter etc. vara av värde även i arbete som ligger utanför det ursprungliga språkområdet.

Det här är inte den lämpligaste platsen för att argumentera för nordiskt samarbete i allmänhet. Mycket kan sägas om den saken. Många är skeptiska till nordiskt samarbete, eftersom det ofta i praktiken stöter på många hinder. Vad man än anser om gemensamma nordiska strävanden på ett mera principiellt plan, kan det helt osentimentalt konstateras, att på det lexikografiska området finns det åtskilliga vinster att göra i det läge som vi befinner oss i, med ständigt ökande kontakt med andra språkområden. De enskilda nordiska ländernas resurser räcker helt enkelt inte till för att tillgodose de önskemål som det inom kort blir angeläget eller nödvändigt att ta hänsyn till. Ju förr vi inser att vi tjänar på att samarbeta och börjar utforma vårt arbete därefter, desto snabbare kan vi dra fördelar av samarbetet. Hade Baldur Jónssons idé redan för tio år sedan omsatts i praktiken, hade vi haft en betydligt bättre beredskap idag.

Antagligen skulle idealet vara att upprätta en eller ett antal gemensamma nordiska databaser med avancerat lexikografiskt innehåll. En sådan åtgärd rymmer många problem – språkliga, kulturella, teoretiska, praktiska etc. men framför allt juridiska, kommersiella, attitydmässiga etc. Det är viktigt att problemen attacker:s snarast, så att en lösning kan

komma till stånd. Som ett steg i den riktningen anordnades en konferens i samarbete mellan Nordiska ministerrådet (genom dess organ Nordisk språksekretariat) och Göteborgs universitet i maj 1990. Den hölls i Göteborg och kunde endast ta emot ett mycket begränsat antal deltagare. Tanken var att dessa skulle väljas så att de representerade olika typer av kompetens och specialiseringar, olika intressenter, olika nationaliteter och språkområden, samtidigt som deltagarna kunde förväntas ha god kontakt med kolleger och intresserade som inte kunde beredas möjlighet att delta. Konferensens beteckning "LEXIKONORD" skulle just utläsas "LEXIKografisk KOMPetens i NORDen", en kompetens som får anses vara betydlig och spridd över så många instanser att endast ett litet urval kunde samlas till den aktuella konferensen. Den probleminventering som blev resultatet av konferensen har publicerats som Rapport 14 från Nordisk språksekretariat och distribuerats generöst över den lexikografiska världen.

Den konferens som genomfördes i Oslo i maj 1991 är ytterligare ett steg i samma riktning, och ett ännu viktigare steg, eftersom denna konferens var öppen för alla intresserade deltagare. Återigen blev det en imponerande manifestation av den samlade lexikografiska kompetensen i Norden. När jag fick det smickrande uppdraget att komma med ett avslutande inlägg vid denna första konferens om lexikografi i Norden, avstod jag medvetet från att referera eller dra slutsatser av bidragen vid konferensen. De publiceras ju härmed. Jag undvek att ge en bred beskrivning av utvecklingen inom ämnesområdet, med litteraturhänvisningar och allt, eftersom det skulle kräva en djupare insikt än den jag förfogar över. För en del allmänna och övergripande men självklart personligt färgade synpunkter på utvecklingen, som kan ha relevans även i det sammanhang det nu gäller, kan jag hänvisa till mitt bidrag i förhandlingarna från Svenskans beskrivning 18 i Uppsala (Ralph 1991). Här har jag inskränkt mig till att förmedla några reflexioner jag gjort under konferensen i Oslo. Ett av de starkaste intrycken är just att det nordiska lexikografiska arbetet för närvarande visar en påtaglig mångsidighet och vitalitet.

Det är ord på väg!

### Litteratur

- Amsler, Robert A. 1980. *The structure of Merriam-Webster Pocket Dictionary*. Austin, Texas. [Dupl.]
- Baldur Jónsson. 1982. Isländskans ställning och roll i det nordiska samarbetet. *Språk i Norden* 1982, s. 5–15.
- Bloomfield, Leonard. 1933. *Language*. London.
- Gavare, Rolf. 1991. Lexikaliskt, datatekniskt stöd vid översättning och författande. *Språkvård* 1/1991, s. 21–27.
- Hannesdóttir, Anna & Ralph, Bo. 1988. Early dictionaries in Sweden: Tradition and influences. *Symposium on Lexicography IV*, ed. by Karl Hyldgaard-Jensen & Arne Zettersten. Tübingen, s. 265–279.
- LEXIKONORD. *Leksikografi i Norden*. 1991. Rapport fra en konferense i Göteborg 11.–13. mai 1990. Oslo. (Nordisk språksekretariats rapporter 14.)
- NFO = *Nusvensk frekvensordbok* 1-4. 1970–80. Stockholm.
- Nordiska rådet, 38:e sessionen 1990. Förtryck B 103/k. Kulturutskottets betänkande över ministerrådsförslag om handlingsprogram för språkligt samarbete i Norden.

- Ralph, Bo. 1991. Strömningar inom ordforskningen. *Svenskans beskrivning* 18. *Förhandlingar vid Artonde sammankomsten för svenskans beskrivning. Uppsala den 25–26 oktober 1990*. Utg. av Mats Thelander et al. Lund, s. 9–28.
- Zgusta, Ladislav, 1988. *Lexicography today. An annotated bibliography of the theory of lexicography*. Tübingen. (*Lexicographica*. Series Maior 18.)



Del 2



Kristina Nikula

## Deskriptiva ordböcker - finns dom?

I detta bidrag konstateras att rubrikens frågeställning egentligen är felaktig beroende på att termen deskriptiv i litteraturen har använts i en alltför vid bemärkelse, då termen beskrivit företeelser på åtminstone två plan, a) forskarens/lexikografens relation till materialet, b) lexikografens relation till ordboksanvändaren. I syfte att skilja de båda aspekterna åt föreslås därför att begreppet deskriptiv används för aspekt a), medan normativ utgör en aspekt av b). Inom b) kan man sålunda skilja mellan normativitet och icke-normativitet, varvid icke-normativitet gäller information om t.ex. bruk, men också om norm.

Det är ett välkänt faktum att få ordboksanvändare läser inledningen till de ordböcker de använder. Bl.a. mångåriga iakttagelser av vilka häpnadsväckande slutsatser språkstudier kan dra på basis av sina ordboksstudier har bekräftat detta och småningom också fått mig att intressera mig för vilken typ av information förord och inledning verkligen erbjuder den som är samvetsgrann och tar del av denna.

En redovisning av syftet är att förvänta i förordet till varje ordbok. Eftersom det ur såväl användar- som författarsynvinkel är viktigt om en ordbok är avsedd att vara deskriptiv eller preskriptiv och att det står klart vad de båda termerna innebär skall jag i det följande kritiskt granska allmänna, standardspråkliga ordböcker i synnerhet med avseende på deras "deskriptivitet", och försöka svara på frågan om det när allt kommer omkring finns ordböcker som är deskriptiva i en eller annan bemärkelse.

För att det skall vara möjligt att besvara frågan diskuteras inledningsvis begreppet deskriptiv mot bakgrund av den diskussion som förts främst inom tysk lexikografisk forskning. Eftersom ordboksanvändaren alltifrån början spelat en central roll diskuteras också i vad mån inledningen ger uppgifter om huruvida utsagorna i ordboken skall uppfattas som normerande eller inte och i vad mån denna information är att lita på. Till grund för undersökningen ligger fyra av de flitigast använda enspråkigt svenska ordböckerna *Illustrerad svensk ordbok* (ISO, 1977), *Bonniers svenska ordbok* (BSO, 1986) och *Svensk ordbok* (SOB, 1986). Trots att *Svenska Akademiens ordlista* (SAOL, 1986) är uttalat normativ utgör den ändå en viktig del av materialet.

### 1. Begreppet deskriptiv

Begreppen deskriptiv och preskriptiv (preskription och norm används i detta arbete som synonymer) uppfattas vanligen som varandras motsats.<sup>1</sup> Om man med en preskriptiv ordbok avser en ordbok där normativa anspråk görs, kan en deskriptiv ordbok enkelt definieras som en ordbok där inga sådana anspråk görs. Det är dock långt ifrån alltid klart vad som egentligen avses med deskriptiv; likaså kan man ifrågasätta om antonymiförhållandet deskriptiv-preskriptiv verkligen gäller.

## 1.1 Deskriptiv i olika betydelser

Att begreppet deskriptiv inom lexikografin används i ett flertal betydelser är ingen nyhet (Mugdan 1985:210, Ripfel 1989:199f.), däremot har veterligen inte ifrågasatts om preskriptiv är den självklara motsatsen till deskriptiv.

Deskriptiv används av vissa forskare (Bergenholtz-Schaeder 1978: 122 ff.) i betydelsen "primär-induktiv". Motsats till induktiv är emellertid inte preskriptiv utan deduktiv. Också deskriptiv och sann, kontrollerbar, används som synonyma begrepp (Mugdan 1985: 210, Wiegand 1986:92). I och med detta ges deskriptiv en betydelse som står i ett motsatsförhållande till värdering, vilket induktiv inte gör, då också vid induktivt förfarande värderingar i viss mån måste företas (Ripfel 1989:200). Saken kompliceras ytterligare då deskriptiv i betydelsen primär-induktiv i vetenskapsteoretisk betydelse ibland (bl.a. av Mugdan 1985:197) indirekt likställs med induktiv i statistisk bemärkelse. Inom induktiv statistik når man visserligen sina resultat på induktiv väg, men för att statistisk metod skall vara tillämpbar ställs bestämda krav på den textkorpus på vilken resultaten bygger. Sammanställningen av korpus blir med andra ord avgörande för om och i vilket avseende resultaten är deskriptiva.

### 1.1.1 Induktiv

Om en ordbok kommer till på deskriptiv i betydelsen induktiv väg i vetenskapsteoretisk bemärkelse innebär det att ordboksförfattaren med utgångspunkt t.ex. i en textkorpus gör utsagor om språkbruket, medan han i det deduktiva fallet utifrån en "färdig uppfattning" om detta söker stöd för sina antaganden i den givna textkorpusen. I vardera fallet kan sålunda en textkorpus komma till användning, men den nyttjas på skilda sätt.

I fråga om ordboksarbete får ett rent deduktivt förfaringsätt betraktas som närmast uteslutet, då det åtminstone vid utarbetande av definitionsordböcker men säkerligen även annars skulle ställa orimliga krav på ordboksförfattaren. Man kan därför utgå ifrån att ordböcker i huvudsak kommer till på induktiv väg med större eller mindre inslag av ett deduktivt förfarande (Ripfel 1989:201).

Det ligger nära till hands att tro att den på sina håll negativa hållningen till normering (jfr Nikula 1991) ligger bakom då ordboksförfattare till varje pris vill karakterisera sin ordbok som deskriptiv i obestämd bemärkelse och att de härigenom strävar till att ge denna en objektiv prägel som ytterligare undestryks genom hänvisning till en bakomliggande textkorpus. För att statistiskt hållbara resultat skall kunna uppnås inom språkvetenskapen krävs emellertid att dessa baseras på en representativ korpus. Inom den lexikografiska forskningen förefaller man emellertid tämligen ense om att det inte är möjligt att skapa en sådan korpus för standardspråk (Gellerstam 1978:17, Mugdan 1985:199, Ripfel 1989:200). Detta beror på att man än så länge inte lyckats avgränsa undersökningsobjektet, standardspråket. Olika försök att definiera standardspråk visar på betydande svårigheter att nå fram till en heltäckande definition (Ammon 1987, Brink 1986). Svårigheten beror främst på att det i detta fall inte är frågan om en texttyp i likhet med t.ex. medicinskt fackspråk, utan att det i stor utsträckning handlar om värderingar av vilken typ av språk som kan accepteras som standardspråk. (Brink 1988:7 ff.) Här är steget inte långt till att man utgår från uppfattningen om en gällande norm, dvs. man arbetar "deduktivt" utifrån föreställningen om en gällande norm, i stället för att, såsom det vore att vänta, arbeta induktivt och med utgångspunkt i korpus försöka nå fram till bruket inom standardspråket och därmed till den norm som gäller för detta. Eftersom det



principiellt inte är möjligt att entydigt avgränsa standardspråk, innebär detta att det på empirisk grund är omöjligt att göra allmängiltiga utsagor om standardspråket (Ripfel 1989:200).

Alla typer av text medför inte samma problem. Det är fullt möjligt att t.ex. vid undersökningen av en författares litterära produktion avgränsa forskningsobjektet, och det torde vara möjligt att i detta identifiera olika texttyper och att dra slumpmässiga stickprov. Eftersom en representativ korpus av samtida standardspråk inte kan skapas, kan man i en ordbok över standardspråket inte heller ge uppgifter som är representativa i vid bemärkelse, utan uppgifterna har en begränsad giltighet för det empiriska material på vilket de baserar sig (Gellerstam 1978:18).

I det fall ordboksförfattaren arbetar med ett material som inte utgör ett representativt stickprov, måste han göra en bedömning av detta med avseende på hur typiskt det är, dvs. beskrivningen måste bygga på värderingar. Risken för felbedömning minskar dock om en omfattande korpus står till ordboksförfattarens förfogande. Vad som avses med en omfattande korpus låter sig emellertid knappast fastställas så länge man inte lyckats avgränsa standardspråkets omfång. En korpus på c. 3 milj. ord bedöms av Ripfel (1989:201) med anledning av ett konkret fall i alla händelser som otillräcklig (jfr Mugdan 1985:203ff.). Statistiska utsagor om standardspråket måste sålunda behandlas med stor försiktighet. Eftersom man dessutom kan beskriva endast mycket begränsade aspekter av språket med kvantitativa metoder kommer jag inte att utförligare behandla statistiska metoder inom lexikologin/lexikografen.

### 1.1.2 Sann, kontrollerbar

En utsaga är deskriptiv om den kan bedömas som sann eller falsk (Wiegand 1987:92). Detta enkla kriterium kan dock inte tillämpas på alla typer av utsagor. En bedömning av om en utsaga är sann eller falsk måste ske på olika sätt beroende på av vilken typ data är. En utsaga av typen a) *Bollen är rund* kan lätt kontrolleras, medan sanningshalten i en utsaga som b) *Bollen är vacker* inte kan fastställas objektivt. Värderande utsagor om språket av samma typ som i b), t.ex. "vulgär", "stötande", liksom i mer eller mindre hög grad bedömningar som "litterär", "provinsiell", "ålderdomlig", "vardaglig" o.s.v., kan inte heller bedömas utifrån sanningsvillkoret och blir inte sanna bara därför att man anför belägg ur en textkorpus och därmed gör dem kontrollerbara intersubjektivt<sup>2</sup> (Jfr Ripfel 1989:201f).

En vanlig uppfattning om vad det innebär att en ordbok är deskriptiv förefaller vara att den redovisar direkt verifierbara data och att den är fri från subjektiva värderingar. Denna föreställning underblåses av företrädare för den korpusorienterade lexikografen, som med hänvisning till en korpus av en viss storlek (som för en lekman på området kan verka enorm) ytterligare understryker verifierbarheten väl medvetna om att en stor del av all den information som ingår i en ordbok utgör resultatet av en tolkning av korpusdata. Inom den korpusorienterade lexikografen uppvisar man sålunda inte sällan en överstor tro på nyttan av en korpus. Det finns emellertid åtskilligt med information som hör hemma i en ordbok som inte utan vidare låter sig direkt härledas ur en textkorpus, t.ex. uppgifter om synonymförhållanden (Wiegand 1986:91), valensförhållanden osv.; inte ens uppgifter om frekvens kan helt mekaniskt härledas ur korpus (Wiegand 1986:92). För att man skall nå fram till uppgifter av denna typ krävs en kombination av korpusdata och ordboksförfattarens kompetens samt ytterligare stöd av substitutionstest, informantutfrågningar etc.<sup>3</sup> (Wiegand 1986:91f). Inte heller ett sådant förfarande gör en utsaga helt kontrollerbar, men minskar i

någon mån subjektiviteten i påståendena (Gellerstam 1987:12). Rent generellt är beskrivningar av språkregler överhuvudtaget inte falsifierbara i begreppets strängaste mening, eftersom de beskriver regler för mänskligt beteende och inte naturföreteelser.

Ordböckerna innehåller sålunda uppgifter som är falsifierbara i en snävare bemärkelse, dvs. intersubjektivt kontrollerbara gentemot en korpus, och sådana som utgör värdepåståenden och bedömningar. Att en ordbok innehåller värderingar och bedömningar betyder dock inte att ordboken automatiskt är att betrakta som normativ (jfr dock Mugdan 1985:210). Visserligen ligger värderingar till grund för normer, men värderingar berör ett vidare fält än normer, som i princip gäller enbart mänskligt handlande. Normbrott kan därtill åtföljas av sanktioner vilket inte är fallet med värderingar. Om värderingar och bedömningar av t.ex. ett materials lämplighet automatiskt ledde till normativitet skulle all språkforskning vara normativ, något som naturligtvis inte är fallet (Ripfel 1989:202).

Deskriptivitet i anslutning till beskrivningar av språket kan sålunda inte förutsätta falsifierbarhet i begreppets stränga bemärkelse annat än i fråga om ytterst triviala iakttagelser som frekvensen av en viss kombination av grafiska tecken, t.ex. grafiska ord, i en avgränsad korpus. Termen deskriptivitet kommer därför inom språkvetenskapen generellt att omfatta påståenden som innehåller värderande element förutsatt att intersubjektiv consensus råder utgående från accepterade eller acceptabla forskningsmetoder.

### 1.1.3 Icke-normativ

Termen deskriptiv i betydelsen induktiv och sann, kontrollerbar anger det sätt på vilket lexikografen väljer att utvinna språkliga data, dvs. med "deskriptiv" anges förhållandet mellan forskaren och undersökningsobjektet utan en tänkt kommunikationspartner. Eftersom förståelsen av en text förutsätter en tolkning inom ett kommunikationssammanhang utgör naturligtvis författaren av belägget/beläggtexterna dock normalt en kommunikationspartner på vägen fram till deskriptionen.<sup>4</sup> Då deskription ställs i motsats till norm (preskription), lyfts begreppet upp på en annan nivå, där det är frågan om hur man väljer att utnyttja de resultat som vunnits på deskriptiv väg. Den verkliga skillnaden mellan norm och deskription består i att norm förutsätter ett normsobjekt (jfr von Wright 1963:70) på vilket normativa krav kan ställas och inte i kontrollerbarheten och sanktionerna, då ju kontrollerbarheten som framgått ovan är ett ganska osäkert kriterium och sanktioner helt och hållet kan utebli vid normbrott. För att en norm skall kunna skapas och genomdrivas krävs därtill en deskription av gällande regler av vilka vissa upphöjs till norm.<sup>5</sup> En deskription ligger sålunda till grund för normering.

Sammanfattningsvis kan konstateras att termen deskriptiv beskriver relationen mellan forskaren och objektet medan termen normativ uttrycker relationen mellan forskaren/lexikografen, ordbokstexten och ordboksanvändaren. Normering förutsätter alltid kommunikation, ett normsobjekt, deskription gör det däremot inte (jfr dock ovan).

#### 1.1.3.1 Intention och funktion

Deskriptiva utsagor om språket är som framgår ovan i 1.1.2 inte fria från värderingar, men så länge de av lexikografen i kommunikationen med ordboksanvändaren är avsedda enbart som information om språkbruket i texter som i och med att de producerats är "historiska" har de tillkommit med en icke-normativ intention. En ordbok som av författaren uppges vara

"deskriptiv" (= icke-preskriptiv, icke-normerande) kan emellertid också utan dennes medverkan få normativ funktion. Uppfattningen av ordboks författaren som en auktoritet och respekten för det tryckta ordet gör att ordboksanvändaren uppfattar utsagor om språkbruket vid en viss historisk tidpunkt x som en uppmaning till ett identiskt språkligt handlande vid en kommande tidpunkt y. Ordboksanvändaren är också benägen att tilldela ordboken normativ funktion så snart dess lämplighet för aktivt bruk vid skapandet av nya texter framhålls. Etiketter av typen "ålderdomlig", "vardaglig" o.d. bidrar likaså lätt till att deskriptiva utsagor utan normativ intention ges normativ tolkning. Så länge en ordbok är avsedd för t.ex. passivt bruk som en nyckel till "historiska" texter är den dock icke-normerande. Detta gäller åtminstone så länge som man inte tar i betraktande andra omständigheter som kan ge upphov till normativ tolkning.

En ordbok som bygger på ett material som vid utarbetandet av ordboken ställts i relation till någon norm och där normvidrigheter uteslutits är till sin intention normerande och man kan här med Ripfel tala om "förtäckt" normativitet, då ordboken skenbart neutralt beskriver ett språk som av ordboks författaren anses vara förebildligt och eftersträvansvärt med tanke på användarna (Ripfel 1989:203). Ett samtida standardspråk beskrivs emellertid inte i ordböcker av denna typ, utan snarare ett rekommenderat normspråk. Beroende på vilken typ av texter som ligger till grund för den förtäckt normativa ordboken måste man sålunda skilja mellan om ordboken är normskapande eller normbevarande. Om textkorpus beskriver normerat språkbruk, kommer ordboken att verka som en förstärkare och upprätthållare av gällande norm, dvs. vara normbevarande, då den däremot inte kan anses vara normskapande, eftersom den inte bidrar till att skapa en ny norm (Ripfel 1989:197, 205).

## 2. Oklarheter i begrepps användningen

Ovan i 1.1.3 framgår att begreppen normativ och deskriptiv inte står i ett motsatsförhållande till varandra eftersom de beskriver företeelser på olika plan. Orsaken till att dessa båda begrepp vanligen uppfattas som varandras motsats är sannolikt det faktum att en deskription normalt ska kommuniceras och att en normering måste utgå från en beskrivning. Detta leder lätt till oklarhet och begreppsförvirring. Till oklarheten bidrar ytterligare att deskriptionen såsom framgår i 1.1.2 också innefattar värderande aspekter.

På frågan om normvidriga stavningar och uttryckssätt med hänvisning till frekvens och autenticitet skall tas med i ordböckerna konstaterar Svensén:

Svaret på den frågan är i praktiken nej. En ordbok för allmänt bruk - dess syfte må vara hur objektivet beskrivande som helst - bär inom sig en stor normerande kraft, i synnerhet när det gäller tecknets uttryckssida. Man kan t.o.m. gå så långt som till att säga att i och med att en viss språklig företeelse utan reservationer tas upp i en ordbok har man i viss mån upphöjt den till norm (Svensén 1987:32).

Det faktum att en språklig företeelse utan reservation upptas i en ordbok innebär dock inte i sig att den upphöjts till norm, men den tolkas sannolikt så åtminstone i fråga om samtida standardspråk. Däremot innebär en reservation<sup>6</sup>, liksom naturligtvis också en rekommendation<sup>7</sup>, normering. Om ordboks författaren, utan att t.ex. i förordet ange detta, undvikit att beskriva sådant språkbruk som han inte uppfattar som lämpligt bara för att ordboksanvändaren inte ska influeras av detta "felaktiga" bruk, har författaren naturligtvis

förtäckt också handlat normerande i sin relation till en tänkt användare, men själva beskrivningen har därmed inte i sig blivit "mindre deskriptiv och mera normerande" av denna anledning.

En liknande sammanblandning av begreppen föreligger uppenbarligen hos Gellerstam då han konstaterar:

Även om man gör en helt deskriptiv ordbok kan man inte undgå frågan om vad som är 'gott språk', den normativa aspekten (Gellerstam 1978:18).

Begreppet det "goda språket" innebär naturligtvis en värdering, men man kan i princip beskriva det man anser vara det goda språket utan att därmed förutsätta att någon annan ska rätta sig därefter. Själva gränsdragningen mellan standardspråk och icke-standardspråk bygger åtminstone delvis på värderingar och delvis på uppfattningar om det "goda språket", eftersom en representativ korpus redan principiellt inte är möjlig, jfr ovan 1.1.1. Även om standardspråket avgränsas utgående från värderingar, dvs. språkkänsla eller språklig "kompetens", betyder detta inte i sig en normering, eftersom språkkänslan är ett uttryck för de regler och normer som styr språklig produktion och interpretation. Frågan om vad som är "gott språk" innebär sålunda principiellt inte med nödvändighet normering. Däremot hyser en ordboksförfattare som använder termen det "goda språket" som beskrivning på något av sina urvalskriterier med största sannolikhet normerande intentioner. En uppgift t.ex. i förordet till en ordbok att den beskriver det "goda språkbruket" kan knappast heller uppfattas som något annat än en rekommendation riktad till ordboksanvändaren, dvs. som en normativ utsaga (Nikula 1990:115).

### 3. Finns deskriptiva ordböcker?

Ovan i 1.1.3 framgår att deskriptiv i litteraturen används i betydelser som till motsats kan ha deduktiv och värderande, men däremot inte som jag försökt påvisa preskriptiv-normativ. Deskriptiv har sålunda använts för att beskriva företeelser på två skilda plan, då termen fått stå dels för forskarens analys och beskrivning av sitt material, dels för hans intentioner vid mötet med ordboksanvändaren.

Så snart en kommunikationspartner kommer med in i bilden, och detta sker automatiskt då ordboksredaktören tar ställning till användarkategori, borde ordboken snarast beskrivas som normerande eller icke-normerande, dvs. klassificeringen borde bygga på förhållandet ordboksförfattare - ordboksanvändare som parter i en kommunikationssituation. Detta innebär att frågan i rubriken, som utgår från uppfattningen om deskriptiv och preskriptiv som varandras motsats, är felaktigt ställd. Ordböcker kan inte, oberoende av hur de står i förhållande till normen, tillkomma på något annat sätt än på basis av beskrivningar av regler, vilka i det normativa fallet antingen förkastas eller upphöjs till norm. Varje ordbok är sålunda i någon mening deskriptiv, men kan ha tillkommit antingen med normativ eller icke-normativ, t.ex. "informerande", intention. En annan sak är sedan att sådana ordböcker som beskriver det samtida standardspråket, oberoende av ordboksförfattarens intentioner av användaren gärna ges normativ funktion.

Förenklat kan tillkomsten av en normerande ordbok betraktas som en tvåstegsprocess där det första steget utgör utarbetandet av en icke-normerande ordbok över vilken i det andra steget läggs ett raster (normen) som släpper igenom det mesta men döljer (förtiger<sup>3</sup> och förser med restriktioner eller rekommendationer) en del.

Gränsen mellan normerande och icke-normerande ordböcker blir lätt oklar och de flesta nuspråkliga ordböcker tenderar att bli normerande då aktiv och passiv funktion vanligen kombineras. Så snart en ordbok är avsedd också för aktivt bruk, kommer normen automatiskt in i bilden, då normspråket i stor utsträckning får antas utgöra modell för böjningen och återspeglas i exempelsatser etc. - om inte skulle ordboksförfattaren förfara bedrägligt mot ordboksanvändaren. Vid tillkomsten av normerande ordböcker går bruket och normen å andra sidan numera i det längsta hand i hand, då man är medveten om att en normering som inte bygger på bruket är dömd att förr eller senare misslyckas (Bartsch 1986:143, Gellerstam 1987:12). Oberoende av om en ordboksförfattare avser att vara normerande eller inte strävar han sålunda först och främst till att klarlägga det allmänna bruket och på normeraren ankommer ytterligare att på grundvalen av sin kännedom om språket som system och språket i bruk ge rekommendationer t.ex. när bruket vacklar. Redan nu tillkommer ju ordböcker av båda sorter på basis av samma material (SAOB och SOB) och de icke-normerande ordböckernas uppslagsord är ofta valda med utgångspunkt i SAOL, på sin höjd med en tillsats av nyord (ISO). Ordboksarbetet hamnar också härigenom ofrivilligt in under inflytande av normen och råkar småningom in i en kretsgång då de texter av vilka bruket avläses kan antas ha tillkommit med beaktande av normen som bygger på bruket.

Ordböckerna har ju hittills i stor utsträckning tillkommit på lexikografens villkor och även om undersökningar gjorts av vad användarna önskar sig av ordböckerna är dessa undersökningar så "ortodoxt" upplagda och genomförda att användarnas verkliga behov och önskemål sällan kommer i dagen (Crystal 1988:78). Eftersom det förefaller tämligen utsiktslöst att ändra ordboksanvändarnas vanor inom överskådlig tid vore kanhända en förändring av ordboken i stället värd att överväga? En förändring i riktning mot öppet i stället för förtäckt normerande ordböcker skulle dock inte behöva innebära en "cementering" av normen (Nikula 1991:248), något som varken är önskvärt eller ens möjligt, utan ordboken skulle kunna utgöra en samling riktlinjer eller rekommendationer (Bartsch 1986: 199f.), som användaren i sista hand accepterar eller förkastar.

### 3.1 Svenska ordböcker och normativiteten

Det torde vid det här laget stå klart att de ordböcker som är icke-normerande eller övervägande icke-normerande när allt kommer omkring är relativt få. Det bästa exemplet på sådana ordböcker förefaller vara ordböcker över äldre språk och dialektordböcker, som väl mera sällan är avsedda för aktivt bruk, medan ordböcker över samtida standardspråk i stor utsträckning är förtäckt normativa. Detta framgår självfallet inte av förordet. I stället är ordböckerna antingen oklara på denna punkt eller visar sig vid närmare granskning inte hålla vad de lovat.

I förordet till ISO sägs syftet vara att utvidga läsarens ordförråd inte minst i fråga om fackuttryck av olika slag. "Illustrerad svensk ordbok har sålunda ett praktiskt syfte, ett folkbildningssyfte" (ISO 1973). Vad detta egentligen innebär tål att tänka på. Som ordboksanvändare föreställer jag mig i vart fall att den vidgade kunskap om språket som ISO avser att bidra till kan vara till praktisk nytta vid *aktivt* bruk, men därom sägas ingenting.

I BSO tas explicit ställning till normativiteten då det i förordet sägs: "BSO är inte, i motsats till Svenska Akademiens ordlista, en normerande ordbok" (BSO 1986). BSO är dock i själva verket den ordbok som kanske allra tydligast föreskriver hur saker och ting i själva verket bör vara (jfr not 6 o. 7).<sup>9</sup>

Också i SOB tas åtminstone indirekt ställning till normativiteten, då utgivaren konstaterar att "Ordboken är i första hand beskrivande" (SOB 1986:V). I och med att denna ordbok också lämpar sig för aktivt bruk har den tagit ett stort steg i normerande riktning. Att denna ordbok också på andra grunder kan anses vara normerande har jag tidigare diskuterat i ett annat sammanhang (Nikula 1990:114ff).

Klarast redovisas syftet i SAOL där det heter: "Svenska Akademiens ordlista är till sin natur en samling rekommendationer. [---] SAOL har från början haft en uttalat språkvårdande inriktning. [---] Denna strävan efter normering har fortsatt genom åren och blivit ett av SAOL:s kännetecken" (SAOL 1986:5f).

#### 4. Kommentar och sammanfattning

I detta bidrag har konstaterats att rubrikens frågeställning egentligen är felaktig beroende på att termen deskriptiv i litteraturen har använts i en alltför vid bemärkelse, dvs. dels i betydelsen analys och beskrivning av de "objektiva fakta" som analysen gett som resultat, dels i betydelsen information om dessa fakta. Termen har sålunda beskrivit företeelser på åtminstone två plan, a) forskarens/lexikografens relation till materialet, b) lexikografens relation till ordboksanvändaren. I syfte att skilja de båda aspekterna åt föreslås därför att begreppet deskriptiv används för aspekt a) medan begreppet normativ (preskriptiv) utgör en aspekt av b). Inom b) kan man sålunda skilja mellan normativitet och icke-normativitet, varvid icke-normativitet gäller information om t.ex. bruk, men också om norm. Vad gäller normativiteten måste vidare skiljas mellan normativ intention och normativ funktion (tolkning). I fråga om deskriptivitet har konstaterats att beskrivningar av språk utan värderande och bedömde synpunkter kan förekomma i fråga endast om mycket begränsade aspekter. Eftersom normering eller preskription inte kan ske utan föregående analys och beskrivning är det svårt att föreställa sig vad som kunde avses med en icke-deskriptiv ordbok.

Eftersom det är uppenbart att ordboksanvändare i allmänhet vid minsta anledning tolkar ordboksutsagorna som normerande, antagligen därför att man söker stöd i en ordbok för att få veta hur man skall uttrycka sig för att uttrycka sig korrekt, faller det sig också naturligt att i förordet karakterisera ordböckerna som normerande eller icke-normerande även om man bortser från de fakta som talar för en sådan terminologi.

Det har också kunnat konstateras att några av de mest nyttjade enspråkigt svenska ordböckerna med undantag av SAOL antingen formulerar sitt syfte tämligen oklart eller i praktiken inte fullföljer angivet syfte.

Eftersom normerande och förtäckt normerande ordböcker låter sig skiljas åt endast i det fall uppgifter ges i inledningen om empirisk bas, sekundärkällor, urvalskriterier för lemmata etc., borde ordboksförfattarna långt mera än vad hittills varit fallet uppmärksamma denna aspekt av normeringen. De ordböcker som uppger att de är icke-normativa borde fylla högt ställda krav på bakgrundsuppgifterna i förord och inledning vad gäller ovannämnda kriterier.

Man kan naturligtvis fråga sig vad en begreppsutredning som den föreliggande har för praktiska konsekvenser. Det förefaller mig för det första självklart att en så väldefinierad begreppsapparat som möjligt är till fördel för allt lexikografiskt arbete - det är alltså här inte frågan om enbart terminologisk precisering. En annan sak är sedan att begreppen norm och deskription utgör endast en delaspekt av lexikografen och att preciseringar på detta område därför inte automatiskt medför några radikala förändringar eller förbättringar av ordböckerna. En så klar och genomtänkt begreppsapparat som möjligt är för det andra grundläggande inom

all undervisning och utbildning där det inte är möjligt att utgå ifrån att vissa förhållanden är självklara eller bekanta från andra sammanhang.

## Noter

1. Så också jag i tidigare arbeten, jfr bl.a. Nikula (1990 o. 1991).
2. Det vore att vänta att åtminstone ordböcker som är jämngamla och bygger på samma korpus skulle sammanfalla med avseende på etiketteringen av uppslagsorden. Då data emellertid är resultatet av en tolkningsprocess finns det t.ex. bland de s.k. runda orden goda exempel på att samma ord värderas olika i ordböckerna ex. *kåt* a. kan uppfattas som stötande (SAOL), starkt vardagligt (SOB), vardagligt (BSO), vulgärt (ISO). Det är här skäl att komma ihåg att ISO är äldst.
3. Om osäkerhetsmoment vid introspektion och informantutfrågning se Mugdan (1985:196).
4. Då man utnyttjar en korpus är det inte frågan om ett mekaniskt utvinnande av data, utan ordboksförfattaren utgör också då han arbetar deskriptivt enligt Wiegand (1986:92) den ena parten i en kommunikationsprocess i vilken han med stöd av sin kompetens formulerar en tolkning av språkbruket.
5. Sådan normeringen genomförs i våra dagar betyder det att det allmänna språkbruket ligger till grund.
6. Goda exempel på reservationer återfinns i BSO i t.ex. följande fall: *vare sig* - Amn. förväxla inte *vare sig* med *varken*. Det heter inte: 'han kan vare sig läsa el. skriva' utan 'han kan varken läsa eller skriva' el. 'han kan inte vare sig läsa el. skriva'; *synvinkel* - *ur* (inte: *från*) *min s. 000*.
7. Inte heller rekommendationer är sällsynta i BSO, ex. *synpunkt* - *från* (hellre än: *ur*) *min s. 000*, *baby* (talspråkets form 'bebisar' brukas inte i skrift; skriv gärna: spädbarn, småbarn).
8. Enligt Malkiel (1989:63) behöver en normativ ordbok inte direkt kritisera för att förhindra ett bruk, då ett förtigande kan verka lika effektivt som kritik.
9. Det anses på sina håll (bl.a. Mugdan 1985:210) sakna betydelse för ordbokens grundhållning om uppgifterna ges i form av beskrivningar eller föreskrifter. En ordbok anses därvid vara icke-normerande så länge uppgifterna är verifierbara.

## Litteratur

- Ammon, U. 1987. Zu den Begriffen 'standardsprachliches Sprachzeichen', 'Standardvarietät' und 'Standardsprache'. I: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung. Band 40/3*, s. 305-316
- Bartsch, R. 1987. *Sprachnormen: Theorie und Praxis*. Niemeyer, Tübingen
- Bergenholtz, H.- Schaefer, B. 1978. Ausblicke auf eine deskriptive Lexikographie. I: *Interdisziplinäres deutsches Wörterbuch in der Diskussion*. s. 116-172. Sprache der Gegenwart Band XLV. Düsseldorf
- Bonniers svenska ordbok*. 1986. 2 uppl. Bonniers. Stockholm
- Brink, L. 1986. Om begreberne rigssprog, standardsprog og dialekt. I: *Nysvenska studier* 66, s. 5-42

- Crystal, D. 1986. The ideal dictionary, lexicographer and user. I: Ilson, R. (ed.). *Lexicography. An emerging international profession*, s. 72-82. Manchester University Press
- Gellerstam, M., 1978. *Välja sina ord. Om ordurval till enspråkiga ordböcker*. Rapporter från Språkdata 9. Göteborg
- Gellerstam, M. 1987. Svenska Akademiens ordlista - elfte upplagan. I: *Språkvård* 2, s. 6-13  
*Illustrerad svensk ordbok*. 1977. 3 reviderade uppl. Natur och Kultur. Stockholm
- Malkiel, Y. 1989. Wörterbücher und Normativität. I: *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Erster Teilband, s. 63-70. Walter de Gruyter. Berlin. New York
- Mugdan, J. 1985. Pläne für ein grammatisches Wörterbuch. Ein Werkstattbericht. I: *Lexikographie und Grammatik. Lexikographica. Series Maior 3*. Max Niemeyer Verlag. Tübingen
- Nikula, K. 1990. Ordböckerna och finlandssvenskan. I: Nikula, K. - Pitkänen, A.J. (red.). *Svenskan i Finland*, s. 109-132. Skrifter utg. av institutionen för filologi II vid Tammerfors universitet, nordisk filologi, Tammerfors
- Nikula, K. 1991. Är norm ett "fult" ord? Om ordböcker och norm. I: *Fackspråk och översättningsteori, VAKKI-seminarium XI*, s. 244-254, Vaasa
- Ripfel, M. 1989. Die normative Wirkung deskriptiver Wörterbücher. I: *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Erster Teilband. s. 189-207. Walter de Gruyter. Berlin. New York
- Svensén, B. 1987. *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. TNC 85. Esselte Studium. Stockholm
- Svensk ordbok*. 1986. Språkdata. Göteborg. Esselte Studium. Stockholm
- Svenska Akademiens ordlista över svenska språket*. 1986. 11 uppl. Norstedts. Stockholm
- Wiegand, H.-E. 1986. Von der Normativität deskriptiver Wörterbücher. Zugleich ein Versuch zur Unterscheidung von Normen und Regeln. I: *Sprachnormen in der Diskussion. Ein Beitrag vorgelegt von Sprachfreunden*, s. 72-101. Walter de Gruyter, Berlin. New York
- Wright von, G.H. 1963. *Norm and Action. A Logical Enquiry*. Routledge & Kegan Paul. London. New York
- Wörterbücher. Ein internationales Handbuch zur Lexikographie*. 1989. Erster Teilband. Hrsg. Hausmann, F.J./Reichmann, O./Wiegand, H.E. /Zgusta, L. Walter de Gruyter. Berlin. New York



Marit Hovdenak

## Norske skuleordlister

Det finst nokså mange rettskrivingsordlister til skulebruk i Noreg. Norsk språkråd står for den offentlige godkjenninga av slike ordlister, og følgjer her eit særskilt regelverk. Eit viktig krav til ordlistene er at dei skal gje fullstendige opplysningar om valfrie former på ein nøytral måte. Sider ved ordlistene som det stundom har stått strid om, er ordutval og uttaleopplysningar. I dei seinare åra har vi fått ordlister tilpassa ulike alderssteg, og dessutan ordlister på edb.

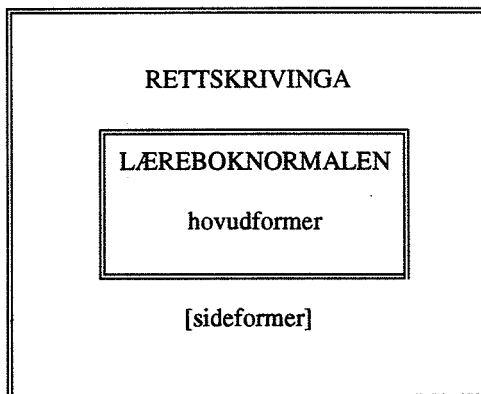
### 1 Innleiing

Det har ofte vore sagt at det norske språket er dårleg dekt med ordbøker, og det kan verke underleg på bakgrunn av den nokså sterke språkinteressa hos nordmenn. Ein type ordbøker har vi likevel eit breitt utval av: *rettskrivingsordlister til skulebruk*. Språksituasjonen i Noreg dei siste hundre åra kan forklare kvifor vi har så mange slike. Det vil denne artikkelen snart komme inn på. Det finst skuleordlister for bokmål og for nynorsk, og for grunnskulen så vel som for den vidaregåande skulen. Til saman blir det temmeleg mange skuleordlister, og dei er meir *like* enn ein kanskje kunne tru. Ordlistene må godkjennast til bruk i skulen. Godkjeningsprosessen er eit av emna for artikkelen.

### 2 Bakgrunn: To nokså like skriftspråk med stor valfridom

Vi har to offisielle skriftvariantar av norsk - *målformene* bokmål og nynorsk. Dei har vore jamstelte skriftspråk i det offentlege livet sidan 1885, og staten har stått for normeringa av dei. I normeringspolitikken dette hundreåret har tilnærming mellom bokmål og nynorsk vore eitt viktig siktemål, andre mål har vore tilnærming til norsk talemål og pedagogiske omsyn. Språkpolitikken har vore gjennomført ved fleire store rettskrivingsreformer. Eit av resultatata er stor *valfridom* i rettskriving og bøyning innanfor kvar av målformene. For kvar målform er det ei trongare norm, *læreboknormalen*, og ei vidare norm, *rettskrivinga*. Læreboknormalen omfattar alle hovudformene, og skal brukast i lærebøker som blir offisielt godkjende til skulebruk og i offentleg teneste elles. Rettskrivinga inneheld heile læreboknormalen og i tillegg sideformer. Sideformene blir òg kalla klammeformer fordi dei står i hakeparentes, skarpe klammer, i ordlistene. Sideformene kan brukast av alle utanom i lærebøker og offentleg teneste. Forholdet mellom læreboknormal og rettskriving kan illustrerast slik:

Offisiell rettskriving innanfor kvar målform:



Etter dei store rettskrivingsreformene dette hundreåret har det vore viktig for styresmaktene å få gjennomført reformene i samfunnet. Her er skulen ein nøkkelinstitusjon. Difor har det vore lagt vekt på å gje ut ajourførte skuleordlister så raskt som råd, og særskilde ordlister til skulebruk har då òg ein hundreårig tradisjon i Noreg. Dei store lærebokforlaga har kvar sine seriar.

Normeringsarbeidet har sidan 1952 gått føre seg i eit statleg organ, Norsk språknemnd, som i 1972 vart avløyst av Norsk språkråd. Språkrådet har (og Språknemnda hadde) rådgjevande status overfor styresmaktene. Språkrådet tek opp aktuelle rettskrivingsspørsmål og gjer vedtak om nye og endra skrivemåtar. (Sjå elles punktet om gransking og godkjenning av ordlister.)

### 3 Korleis ser skuleordlistene ut? - ein presentasjon

Ei typisk skuleordliste er ei alfabetisk liste over det sentrale ordtilfanget i anten bokmål eller nynorsk, utgjeven i hefte- eller bokform. Ordtilfanget kan variere frå 10 000 til 35 000 ord (sjå elles punktet om omfanget seinare). Skrivemåte og bøyning av orda er faste opplysningar. Ofte blir det gjeve ein del uttaleopplysningar og tydingar til orda. Før var ordlistene nokså snaue med lite ordforklaringar. I nyare ordlister er det gjerne meir synonym og andre forklaringar. Somme av desse nyare heiter difor *ordbok*, ikkje *ordliste*.

Utsnitte frå nokre bokmålsordlister nedanfor viser korleis notasjonen og mengda av opplysningar kan variere:

døpe, -te  
døpefont, -en  
dør, -a/-en

(Krogslrud og Seip)

døpe, -te  
døpefont, -en  
dør, -a/-en

(Bakke og Dalene)

døpe, -te, -t  
døpefont/en  
dør/a el -en

(Torvik og Gundersen)

døpe -te har -t navnsatte døpenavn  
dør -a/-en dørhåndtak

(Knappen og Thoresen 1988)

Utsnitta nedanfor er frå to nynorskordlister. Den første er ei nyare ordliste med e-infinitiv, den andre er ei eldre med a-infinitiv.

ofre -a  
ofring/a, -ar  
ofse -a V  
ofse/n  
ofseleg

(Bjones og Dalene)

læra, -te  
lærar/en  
lærarinn/e, -a  
lærd, -  
lærdom/(m)en  
lær/e, -a S

(Eskeland og Torvik)

Utanom sjølve opplistinga av orda inneheld ei ordliste føreord og ei *retteiing* om korleis lista skal brukast, pedagogisk tilrettelagd for elevar. Denne retteiinga kan vere nokså fyldig, som ein liten grammatikk.

#### 4 Godkjenningsordning og ordlistereglar

Noreg har ei offentleg godkjenningsordning for lærebøker i grunnskulen og den vidaregåande skulen. Ordninga sorterer under Kyrkje-, utdannings- og forskingsdepartementet. Alle lærebøker som blir brukte i dei to skuleslaga, skal vere godkjende. Kravet om godkjenning gjeld likevel berre dei sentrale lærebøkene i skulen, ikkje hjelpebøker som arbeidsbøker, emnehefte og oppgåvesamlingar. *Ordlister* er her eit unntak, dei er godkjenningspliktige. Bøkene får godkjenning for fem år om gongen.

Formålet med godkjenningsordninga er å sikre ein høg fagleg og språkleg standard i lærebøkene. Språkrådet granskar språkbruken i dei godkjenningspliktige lærebøkene som ledd i godkjenninga. Skuleordlistene er det Språkrådet som har godkjenningsoppgåva for åleine, bortsett frå at Grunnskolerådet kan komme med ei fråsegn om pedagogiske sider ved nye ordlister for grunnskulen.

Til grunn for arbeidet med godkjenning av ordlister ligg det *Reglar for utforming av ordlister til skulebruk*, utarbeidde av Norsk språknemnd. Reglane vart godkjende av Kyrkje- og undervisningsdepartementet og er daterte 17. april 1959. Dei har ikkje vore reviderte sidan, men det har innarbeidd seg ein praksis for somme spørsmål dei ikkje dekkjer, til dømes for omfanget av ordlistene.

## 5 Omfanget av ordlistene

Reglane for utforming av ordlister seier at "Det sentrale ordfanget i kvart av måla må vera med. Elles kan ordlisteforfattaren i ein viss mon avgjera kva ord han vil ha med." Noka nærare avgrensing av kva som skal reknast med til det sentrale ordtilfanget, har vi ikkje.

Ein ting som står fast, er at kvart oppslagsord skal oppførast med alle valfrie former det har. Dei fleste orda har rett nok berre éi form, men nokså mange har to eller fleire former. Til dømes har adjektivet *gammal* éi hovudform og éi sideform i nynorsk og to jamstelte hovudformer i bokmål:

nyno. **gammal** [gamal]  
 bm. **gammal** el. **gammel**

Om dei ymse formene er mykje eller lite brukte, spelar inga rolle for oppføringa. Skuleordlistene skal alltid gje fullstendige opplysningar om valfridommen i den offisielle rettskrivinga. (For unntak sjå punkta om oppslagsord og bøyingsopplysningar.)

Noko allment kan seiast om kva som høyrer med til det sentrale ordtilfanget: For det første er det opplagt at vanlege ord i dei *lukka* ordklassane - infinitivsmarket, artikkelane, pronomena, preposisjonane, konjunksjonane og interjeksjonane - skal vere med. For det andre må ei mengd ord frå dei *opne* ordklassane substantiv, verb, adjektiv og adverb med. For det tredje må ordlistene innehalde ei viss mengd grunnord; samansette og avleide ord blir det liten plass til i små ordlister. For det fjerde: Ordlistene byggjer i stor grad på eldre ordlister og ordbøker og til dømes på innstillingar frå rettskrivingskomiteane. Står eit ord først i ei kjend ordliste, får det ein viss autoritet som gjer det vanskeleg å stryke ordet, sjølv om det skulle verke uaktuelt elles. For det femte følgjer ordlistene den offisielle språkpolitikken i det å vere atterhalden med å ta inn nye lånord. (Sjå elles det neste punktet, om ordutvalet i nynorske ordlister.)

Ordlistene plar dessutan ha med ein del oppslagsord som ikkje høyrer med til det eigentlege ordtilfanget - forkortingar og stadnamn, særleg namn med vanskeleg skrivemåte e.l.

Dei ordlistene som i seinare tid er laga særskilt for barneskulen, har eit omfang på om lag 10 000 ord. Dei største ordlistene for høgare skulesteg har med godt over 30 000 ord. Språkrådet meiner at ordlister for den vidaregåande skulen helst bør ha med meir enn 20 000 ord.

## 6 Ordutvalet i nynorske ordlister

Spørsmålet om kva for nokre ord som skal stå i ordlistene, har vore eit stridsspørsmål i nynorsk. Striden har mykje dreidd seg om kvar grensa mot bokmål skal gå, og konkret om danske og tyske lånord, ikkje minst dei lett kjennelege "anbeheitlese"-orda.

Det har skorta på store ordbøker som kunne gje opplysningar om ordtilfanget i moderne nynorsk. Den førande ordlista har vore Samlags-ordlista, som Alf Hellevik har stått for i ein mannsalder. Ordutvalet i den har vore rekna som temmeleg absolutte rammer for gode nynorske ord. Denne ordlista har med eit såkalla fornorskingstillegg, ei liste over ein del bokmålsord som i regelen ikkje blir brukte i nynorsk, eller har ei anna form, med opplysning om tilsvarande nynorske ord og former.

Spørsmålet om ordtilfanget i nynorsk har vore eit problem i skulen, der elevar og lærarar ofte har vore i tvil om dei og dei orda kunne brukast når dei "ikkje stod i Hellevik". Språkrådet sette i 1981 ned ei nemnd til å greie ut om ordtilfanget i nynorsk. Resultatet vart ei varsam opning for dei orda som medvite har vore haldne ute frå nynorsk. Dessutan vart det i eit rundskriv opna for ein meir liberal rettepraksis i skulen. Det vart òg vedteke ei ordning for godkjenning av nye ord i skuleordlistene. Fagnemnda i Språkrådet skal vurdere omstridde ord som ein ordlisteredaktør ønskjer å ha med. Kva nye ord som kan takast med i ordlistene, blir kunngjort i årsmeldingane frå Språkrådet.

## 7 Oppslagsorda

Oppslagsord som har to valfrie former, må i regelen først opp to stader i ordlista. Det gjeld til dømes adjektivet

nyno. **aleine** el. **åleine**  
bm. **aleine** el. **alene**

I nynorskordlister må ordet stå både under *a* og under *å*, i bokmålsordlister må *ale(i)ne* stå ein eller to stader etter om det er avstand mellom dei to formene. I små ordlister er det gjerne ingen ord i alfabetet mellom *aleine* og *alene*:

*ale(i)ne ensom, for seg selv, forlatt, enslig*

(Knappen og Toresen 1988)

al[e]ne el al[å]ne  
aleksandr[er]ner/en, -e[r]  
alemann[er]en, -e[r]  
alen/en el alna, fl alen  
el alner  
alene el al[e]ne

(Berulfsen og Lundebø)

I ordlistereglane heiter det: "Reint allment gjeld at lista skal gje fullstendig og heilt nøytral opplysning om valfridom i ordformer og bøyingsformer. Skipnaden skal for vanlege skuleordlister vera strengt alfabetisk. Det gjeld også avleiings- og bøyingsendingar. Både i ordlistene og i undervisninga må det gjerast klart at dette ikkje tyder noka vurdering av dei einskilde formene." Dette prinsippet skal sikre at brukarane skal kunne velje fritt mellom alle valfrie former, og at språkpolitiske synspunkt hos ordlisteredaktøren ikkje skal kunne påverke oppstillinga.

Likevel gjev reglane høve til å opplyse om visse former, særleg sideformer til heile kategoriar, berre i rettleiinga. Ordlisteredaktørane har nytta dette høvet langt på veg. I praksis blir det då ein del former som ikkje står på sin alfabetiske plass i ordlista, og som av den grunn både blir lite kjende og gjerne rekna som perifere.

Sideformene blir alltid markerte grafisk med skarpe klammer, og dessutan kan hovudformene prentast med sterkare skrift enn sideformene. Ofte blir bøyingsopplysningar om eit ord berre gjevne der hovudforma står:

[ålmenn] sjå **allmenn**

(Bjones og Dalene)

I nynorsk har infinitiv av verb valfritt endinga *-a* eller *-e*. Her går regelen ut på at ordlisteredaktøren vel ei av endingane og gjennomfører den for alle verb i ordlista, mens det i rettleiinga skal stå at *-a* og *-e* er valfrie. A-infinitiv har no gått meir og meir av bruk i ordlistene. I begge målformene er kløyvd infinitiv sideform, men i bokmålsordlistene treng ein ikkje opplyse om denne sideforma i det heile. I nynorsk er det òg fleire andre kategoriar av sideformer som det er nok å opplyse om i rettleiinga, døme: *ei jenta, vanlig, å minnas*.

## 8 Bøyingsopplysningar

Ordlistene opplyser oftast om bøyinga av orda ved å føre opp bøyingsendingar. Bøyingsendingane viser samstundes ordklassen. Uregelrette ord må det gjevast meir fullstendige opplysningar om. For regelrette substantiv er oftast berre bunden form eintal (og dermed genus) vist. Rettleiinga forklarar då korleis den fullstendige bøyinga skal vere. Som for oppføringa av oppslagsorda gjeld det at bøyingsopplysningane skal gjevast på ein nøytral måte utan preferansar. Også bøyingsendingar skal stå alfabetisk (sjølv om det finst unntak):

bm. **kaste** -a/-et  
**dør** -a/-en

Nokre valfrie bøyingsformer treng ein ikkje opplyse om ved kvart ord, det er nok å omtale dei i rettleiinga. Det gjeld i bokmål substantiv på *-het*, som har valfritt *-a* eller *-en* i bunden form eintal, men som kan stå med berre *-en* i sjølve lista. I nynorsk gjeld det sideformer i bøyinga av både substantiv, adjektiv og verb.

For barneskulen har det dei seinare åra komme fleire ordlister med fullstendige bøyingsformer, ikkje berre endingar, også for heilt regelrette ord. Desse ordlistene har eit heilt anna preg enn dei tidlegare:

bås, båsen, båser, båsene  
båt, båten, båter, båtene  
båtnaust, båtnaustet, båtnaust, båtnausta  
el. båtnaustene

(Cordtsen og Sandvei)

## 9 Uttaleopplysningar

Reglane for utforming av skuleordlister seier ingen ting om uttaleopplysningar, og ordlistene har jamt over sparsamt med slike. For *heimlege* ord har det først og fremst i nynorsk vore vanleg å opplyse om vokalkvaliteten (særleg av *o*) med aksentar, men for det meste berre for å skilje homonym (accent grave for open, accent aigu for trong uttale):

hóljen (låg haug)  
hòljet (knapp-)

(Hellevik)

Ved *framandord* som ikkje følgjer norske uttalereglar, blir det i mange ordlister gjeve uttaleopplysningar i ein parentes, det gjeld særleg ord frå engelsk og fransk.

Trykkmarkering er mykje brukt. Somme ordlister markerer òg tonem saman med trykket eller for seg. Det har vore vanleg å markere hovudtrykket i ein del ord som ikkje har førsteleddstrykk i "standard austnorsk", det gjeld særleg ord av romansk opphav som *kontor* og *stasjon*. Tidleg i 1980-åra var det mykje diskutert både i og utanfor Språkrådet om ordlister burde ha med trykkmarkeringar. Sidan vi ikkje har offisielle uttalereglar i norsk, meinte mange at då skal heller ikkje ordlister velje ut den uttalen som har størst prestisje og opplyse om den. I denne saka var det meiningsskilnad mellom bokmålsseksjonen og nynorskseksjonen i Språkrådet, men resultatet vart til slutt det same for begge måla: I ordlister som har trykkmarkeringar, skal det forklarast i rettleiinga at desse opplysningane ikkje er normative. Trykket blir markert med prikk under eller med utheva skrift:

okørjen (-gul) (Hellevik)  
okkult (løynteg,  
mystisk); -smejn

effekt, -en (virkning) (Bakke og Dalene)  
effektiv, -t (virkningsfull)

## 10 Ordforklaringar

Tradisjonelt har skuleordlistene vore knappe i utforminga, med liten plass til anna enn formelle opplysningar. Opplysningar om tyding har vore noko brukte for å skilje mellom homonym, men òg til ein viss grad for å forklare vanskelege framandord. Ordlisteredaktørane fører ikkje sjeldan opp norske avlørsarar som dei ønskjer å framheve framfor lånord. Denne linja er særleg følgd i nynorskordlister, men òg i bokmålsordlister, særleg før. Til dømes står adjektivet *berømt* med 'namngjeten' som synonym i Helleviks ordliste, mens det ikkje er ført opp synonym til oppslagsorda *namngjeten* og *vidgjeten*.

I nyare ordlister er det ført opp fleire synonym og andre ordforklaringar enn før. Utsnittet nedanfor er frå ei eldre ordliste som la vekt på å få med ordtydingar:

døpe -tø || *kristne, øse vann, innvie* (Bøe og Fossestøl)  
døpefont -en, -navn  
dør -a

Det neste utsnittet er frå ei heilt ny ordliste:

døpe -te (Kulbrandstad og Lundeby)  
døpefont -en (kum til å ha  
døpevann i)  
dør -a/-en (-feie for sin egen d--:  
ordne sine egne saker først)

Helleviks ordliste har med mange slags opplysningar i ei komprimert form:

odaliskjen (Hellevik)  
(haremsslavinne)  
oddjen (spyd-)  
odde/n (nes)  
odde adj (o. ei par)  
oddetaljet  
offelde (utt. åffsaid)  
ofre -a; ofring/a, -er  
ofse -a (v)  
ofse/n; -leg -  
ofsen -en, -e ei -i  
[-ent], ofse

To nyare ordlister for barnesteget, med fulle bøyingsformer, har med ein del tydingsopplysningar, men på nokså ulik måte. Den første har alle opplysningane i ein parentes etter bøyingsformene. Den siste gjennomfører synonym i kursiv og andre slags tydingsopplysningar i vanleg skrift:

bås båsen båsar båsane (kua står på båsen)  
 båt båten båtar båtane (robåt, motorbåt)  
 båtstø båtstøa båtstøer båtstøene (båt plass) (Hellevik og Hjulstad)

bås båsen fl båsar båsane rom for husdyr  
 i fjøs  
 båt båten fl båtar båtane farty, skip  
 båtshake båtshaken fl båtshakar  
 båtshakane (Knappen, Thoresen og Ølmheim)

## 11 Ei sam-norsk ordliste

Det finst ei ordbok - eller ordliste - som fører opp bokmålsord og nynorskord saman i éi liste. Det er *Norsk ordbok* av Dag Gundersen. Ord som berre er godkjende eller brukte i bokmål, har Gundersen markert med +, ord som berre er godkjende eller brukte i nynorsk, er markerte med \*. Ord som er godkjende i begge målformene, er umarkerte. Ordboka har lite ordforklaringar og ingen uttaleopplysningar:

\*eg jfr \*jeg  
 egal A  
 egalitet|en  
 \*eggvarmar|en,  
 \*eggvarmer|en (Gundersen)

Denne ordboka vart aldri godkjend til skulebruk. Språknemnda, og seinare Språkrådet, meinte boka ville bli for komplisert å bruke for elevane, og dessutan spela det inn at det var så mange uklære punkt når det gjaldt ordtilfanget, der trongst det ein grenseoppgang mellom bokmål og nynorsk. - Om boka ikkje vart godkjend, selde ho godt for det. Det bør nemnast at Språkrådet seinare vedtok å opne for denne typen sams ordlister for dei to målformene.

## 12 Gransking og godkjenning av ordlister

Arbeidet med å granske ordlister er nokså spesielt og svært tidkrevjande. Det er eit nitid arbeid og krev ein oppmerksom lesar. Arbeidet blir gjort av sekretariatet i Språkrådet og av og til av eksterne konsulentar. Det går først og fremst ut på å kontrollere at opplysningane om skrivemåte og bøyning er korrekte og fullstendige etter gjeldande rettskriving, og at oppføringsmåten er konsekvent. Somme tvilsspørsmål som kjem opp når ein granskar ordlister, blir lagde fram for fagnemnda i Språkrådet. Granskinga blir særleg tidkrevjande når det gjeld heilt nye ordlister. Ei ordliste blir godkjend når Språkrådet har gjennomgått henne og funne at ho stettar dei krava som blir stilte, og når Grunnskolerådet eventuelt har komme med ei fråsegn om opplegget for ordlista. Når godkjeningsperioden på fem år nærmar seg slutten for ei ordliste, må forlaget søkje om godkjenning for ei ny utgåve eller nytt opplag. Ordlistene må reviderast i pakt med nye rettskrivingsvedtak.

Arbeidsmengda med ordlister varierer sterkt frå år til år. Etter den store rettskrivingsrevisjonen for bokmål i 1981 vart dei fleste bokmålsordlistene fornya, med påfølgjande arbeidspress for Språkrådet. No har Språkrådet innført ei ny ordning for behandling av rettskrivingssaker. Dei skal vurderast samla kvart fjerde år, slik at ein i normeringa kan sjå



endingane i ein større samanheng. Denne fireårsordninga blir utprøvd i januar 1992. Etter det ventar vi at forlaga reviderer ordlistene sine.

### 13 Språkrådet ordlisteredaktør?

Språknemnda og Språkrådet har ikkje til no sjølv gjeve ut skuleordlister eller andre rettskrivingsordlister, men det har vore planlagt ved fleire høve.

Då Språknemnda hadde fått læreboknormalen av 1959 i hamn, ville nemnda få utarbeidd rettskrivingsordbøker for dei to måla. Dei var tenkte å vere mønsterordbøker for mindre ordlister til skulebruk, og skulle etter planane ha om lag same omfang som Matias Skard: *Nynorsk ordbok* eller *Retskrivningsordbogen* i Danmark. Desse mønsterordbøkene skulle vise dei nye reglane i læreboknormalen i praksis og elles regulere mange detaljar som ikkje var særskilt nemnde der. Av ymse grunnar vart ikkje desse planane fullførte.

Noko seinare, ved midten av 1960-åra, kom det i gang arbeid med ein norsk ordregistrant ved Universitetet i Bergen. Ordregistranten er eit datamaskinelt register over det sentrale allmennspråkelege ordtilfanget i bokmål og nynorsk. Språkrådet og Norsk leksikografisk institutt kom snart inn i arbeidet med registranten. Tanken var mellom anna å gje ut ordbøker på grunnlag av registranten, men arbeidet med å ajourføre registranten stoppa opp, og registranten har vorte lite utnytta.

Eit tiltak som derimot fekk ein lykkeleg utgang, var arbeidet med handordbøkene: *Nynorskordboka* og *Bokmålsordboka* kom ut i 1986. Sidan den tid er det utarbeidd ei større rettskrivingsordliste på grunnlag av Bokmålsordboka, og den vil liggje føre i 1992. Arbeidet med ei tilsvarende ordliste på grunnlag av Nynorskordboka har ikkje kunna komme i gang.

### 14 Ordlister på edb

Elektroniske ordlister blir meir og meir vanlege, ikkje minst som ordrettingsprogram i tekstbehandlingssystem. Ordrettingsprogram er ei god hjelp når ein vil kontrollere om orda er rett stava i ei tekst i datamaskinen. Dataordlister har ei viktig språkpolitisk side ved at dei kan verke sterkt normerende på skriftspråka. Vi ser tendensar til at det kjem ordlister med langt trongare språknorm enn offisiell rettskriving, både fordi mange brukarar ønskjer ei trong norm, og fordi valfrie former tek stor plass.

Det er funne nokså mykje feil og avvik frå offisiell rettskriving i ein del av dataordlistene. Det er uheldig at vi ikkje kan stole på at dei viser korrekt språk. Ordlister som skal brukast i skulen, må ha med dei valfrie formene, også sideformer. Ordlister for offentleg bruk elles må vise dei valfrie formene i læreboknormalen.

Språkrådet har hevda overfor Kyrkje-, undervisnings- og forskingsdepartementet at ordlister i ordrettingsprogram til bruk i skuleverket bør godkjennast på linje med vanlege trykte ordlister, og med Språkrådet som godkjenningsinstans. Departementet har til no vore negativt til dette.

### 15 Ordlistene og brukarane

Vi har lite systematisk kunnskap om korleis elevar bruker ordlister, og om korleis ordlistene bør utformast for å vere mest mogleg til nytte for brukarane. Skuleordlistene er laga først og

fremst for at elevar skal stave og bøye orda rett. For å finne eit ord i ei ordliste må ein ha ein idé om korleis det blir stava, og dessutan kunne alfabetet. Dei vanlege orda som alle skriv rett, som *i*, *dag* og *på*, har sikker plass i ordlista, sjølv om ingen slår opp på dei. Derimot er det vanskeleg å finne eit spesielt ord som *becquerel* dersom ein trur det blir stava *bekke-* eller *beque-* (om det då i det heile står i lista). Ordlistebrukarar må òg ha ei viss grammatisk innsikt for å resonnerer seg fram til oppslagsforma av eit ord og for å bøye det rett ut frå endingane. I vanlege ordlister er det lite med tydingsopplysningar som kan vere til hjelp.

Ei ordliste er altså eit krevjande hjelpemiddel. Vi har mange vitnemål om at også lærarar og andre vaksne brukarar kan komme til å tolke ordlista gale, mellom anna fordi dei ikkje kjenner føreord og rettleiing. Berre det å bøye eit regelrett substantiv i fleirtal kan vere for vanskeleg, iallfall i nynorsk. Det er såleis ikkje utan grunn at forlaga begynte å lage ordlister med fulle bøyingsformer for barnesteget, slik det er nemnt tidlegare. Det var Folkeskolerådet (det som seinare vart til Grunnskolerådet) som i 1970-åra rådde til å utarbeide slike ordlister. Språkrådet har fleire gonger teke til orde for at ordlistene bør vere tilpassa alderssteget.

Etter mønsterplanen for grunnskulen skal elevane tidleg lære å bruke ordliste. Norskklærebøkene har øvingsoppgåver, til dømes med å ordne ord alfabetisk, å finne synonym og å skilje mellom homonym. Opplæringa i ordlistebruk varierer mykje frå lærar til lærar. Det varierer òg kor gamle elevane er når dei får utdelt ordlister, det kan til dømes vere i fjerde eller sjette klasse. (I den vidaregåande skulen kjøper elevane bøkene sjølve.) Det viser seg ofte at det er dei flinke elevane som gjer seg mest nytte av ordlister.

Nokså nyleg er det vorte tillate for elevane å bruke ordliste eller ordbok ved eksamen i norsk, frå våren 1991 i grunnskulen, frå våren 1992 i den vidaregåande skulen. Som grunn for å gje løyve til å bruke ordliste blir det vist til at elevane skal få øving i å bruke hjelpemiddel gjennom skuleåret, og då bør det ikkje vere annleis til eksamen. Kanskje kan ordninga føre til at det blir lagt mindre vekt på å lære rettskriving, og at elevane blir usikre så dei bruker for mykje tid ved eksamensbordet på å slå opp i ordlista.

Dei valfrie formene fyller mykje opp i ordlistene og gjer dei meir innfløkte enn dei elles ville vere. På den andre sida har alle stor sjanse for å finne ei form som samsvarer godt med talemålet. Det er like fullt eit faktum at mange av dei valfrie formene er lite utnytta. Særleg dei sideformene som berre er nemnde i rettleiingane, er lite brukte og kjende. Ein kan diskutere om den store valfridommen er mest til vinning eller tap for språkbrukarane. For eigen del vil eg seie at han er så stort eit gode at vi heller får dragast med ein del ordlistefyll.

Eit ønske for framtida er at vi får *barneordbøker* som legg like stor vekt på tydinga og bruksmåtane til orda som på formelle ting, og som gjev borna lett tilgjengeleg kunnskap om kvardagsnorsken (og helgenorsken med, for den del). Både danskane og islendingane har fått fine, illustrerte barneordbøker nyleg. Ein kan spørje seg kvifor vi ikkje har slike i Noreg. Måljamstellinga, valfridommen og stadige rettskrivingsendringar som snart gjer ordlistene uaktuelle, er nok mykje av svaret. Men også norske skuleelevar fortener hjelpemiddel som er meir interessevekkjande enn dei tradisjonelle ordlistene.

## 16 Sluttord om Språkrådet og ordlistene

Normering av skriftspråka er ei hovudoppgåve for Språkrådet. Å gje ut eigne rettskrivingsordlister skulle vere ein høveleg måte å kunngjere resultatata av normeringsarbeidet på, men først med den nye bokmålsordlista kjem det ei ordliste frå Språkrådet. Språkrådet bør no òg gripe sjansen til å få utarbeidd dataordlister for begge målformene, pålitelege mønsterordlister.

Trengst ordlistegranskinga? Det ser i alle fall ut til at norsk rettskriving er såpass innfløkt at det må spesialistar til for å avgjere somme tvilsspørsmål. Eit rettskrivingsvedtak kan ha følgjer ein ikkje ser med det same (til dømes gje nye inkonsekvensar). Difor er det nok nødvendig med ein viss kontroll for at ordlistene verkeleg skal vise den offisielle rettskrivinga, men granskinga kan kanskje forenklast noko.

Ordninga med å godkjenne nye ord som kan takast med i nynorske ordlister, er ingen god situasjon for nynorsken, slik eg ser det. Ei forbodslinje er sjølvsagt den enklaste måten når ein vil opplyse om kva som er god eller dårleg nynorsk, men kanskje kan ho skade nynorsken meir enn ho stør.

Reglane for utforming av skuleordlister kunne gjerne vore ajourførte. Det bør likevel ikkje gjerast slik at oppsettet i ordlistene blir meir standardisert. Det er mange måtar å presentere ordkunnskapen på, og dei må komme elevane til gode.

### Skuleordlister det er vist til i artikkelen

- Bakke, Karl og Halvor Dalene. 1982. *Ordliste for grunnskolen. Ny utgave 1982 ved Kåre Fjørtoft*. Oslo: Cappelen.
- Berulfsen, Bjarne og Einar Lundeby. 1984. *Aschehougs ordlister. Bokmål. 6. reviderte utgave*. Bokmål. Oslo: Aschehoug.
- Bjones, Jon og Halvor Dalene. 1985. *Nynorsk ordliste for alle*. Oslo: Universitetsforlaget. (3. opplag)
- Bøe, Lars og Bernt Fossestøl. 1968. *Skuleordbok med synonymmer*. Oslo. Johan Grundt Tanum Forlag. (3. opplag)
- Cordtsen, Else og Marius Sandvei. 1988. *Ordliste for grunnskolen. 3. utgave*. Oslo: Gyldendal Norsk Forlag.
- Eskeland, Ivar og Ingvald Torvik. 1961. *Aschehougs ordlister. Nynorsk. Fullstendig liste, alle former oppførte. 2. utgåva*. Oslo.
- Gundersen, Dag. 1966. *Norsk ordbok. Bokmål og nynorsk*. Oslo: Universitetsforlaget
- Hellevik, Alf. 1990. *Nynorsk ordliste. Større utgåve med fotnotar, fornorskingsstillegg og liste over forkortingar. 6. utgåva*. Oslo: Det Norske Samlaget. (2. opplag)
- Hellevik, Alf og Håvard Hjulstad. 1983. *Ordliste for grunnskulen med ordforklaringar, synonym og fulle bøyingsformer*. Oslo: Det Norske Samlaget.
- Knappen, A.E. og A.Thoresen. 1983. *Ordliste for ungdom med synonymmer, ordforklaringar og eksempler fra ordenes bruksområder. Bokmål*. Oslo: Aschehoug.
- Knappen, A.E. og A. Thoresen. 1988. *Ordliste for barneskolen med synonymmer, ordforklaringar og eksempler fra ordenes bruksområder. Bokmål*. Oslo: Aschehoug. (5. opplag)
- Knappen, A.E., A. Thoresen og P.A. Ølmheim. 1980. *Ordliste for barneskulen med synonym, ordforklaringar og døme frå bruksområda til orda*. Nynorsk. Oslo: Aschehoug.
- Krogstad og Seip. 1982. *Norsk rettskrivningsordliste. Ny utgave 1982 ved Finn-Erik Vinje*. Oslo: Cappelen.
- Kulbrandstad, Lars Anders og Einar Lundeby. 1989. *Bokmålsordliste for den videregående skolen*. Oslo: Norsk Undervisningsforlag.
- Torvik, Ingvald og Dag Gundersen. 1984. *Ordliste for alle. Bokmål. 5. utgave ved Dag Gundersen*. Oslo: Universitetsforlaget.

**Annan litteratur**

- Almenningen, Olaf, Thore A. Roksvold, Helge Sandøy og Lars S. Vikør (red.). 1985. *Språk og samfunn gjennom tusen år*. Oslo.
- Bokmålsordboka*. 1986. Oslo.
- Dansk Sprognævn. 1986. *Retskrivningsordbogen*. København.
- Haugen, Einar. 1969. *Riksspråk og folkemål. Norsk språkpolitikk i det 20. århundre*. Oslo.
- Haugen, Einar. 1984. Introduction. I: Eva L. Haugen. *A Bibliography of Scandinavian Dictionaries: 1-61*. New York.
- Hellevik, Alf. 1991. Ordlista som norm og som rettleiing. I: *Norsk læreren* nr. 3 1991.
- Hellevik, Alf og Einar Lundeby (red.) 1964. *Skriftspråk i utvikling. Tiårsskrift for Norsk språknemnd 1952-1962*. Oslo.
- Jahr, Ernst Håkon. 1979. Om "hovedformer", "jamstilte former" og "sideformer" i skriftnormalane. I: Vikør, Lars S. og Geirr Wiggen (red.). *Språklig samling på folkemåls grunn. Artikler frå bladet Språklig Samling*. Oslo.
- Lauvhjell, Arne (red.). 1983. *Heit strid om nynorsk. Dokument og meiningar 1980-83*. Oslo.
- Lundeby, Einar, Leif Mæhle og Oddrun Grønvik (red.). 1988. *Fornyng og tradisjon. Språkvern og språkrøkt 1972-1988*. Oslo.
- Norsk språknemnd. *Årsmelding*. 1952-71.
- Norsk språkråd. *Årsmelding*. 1973-.
- Nynorskordboka*. 1986. Oslo.
- Petersen, Pia Riber. 1987. *Politikens Børneordbog*. København.
- Reglar for utforming av ordlister til skulebruk*. Norsk språknemnd 17. april 1959.
- Regler for godkjenning av lærebøker for grunnskole og videregående skole*. Fastsatt ved kongelig resolusjon av 13. januar 1984.
- Sigurður Jónsson. 1988. *Barnaorðabókin*. Reykjavik.
- Skadberg, Kåre. 1988. *Lærebøkene - ein føresetnad for eit høgt utdanningsnivå*. I: Lundeby o.a. (red.)
- Skard, Matias. 1965. *Nynorsk ordbok for rettskriving og litteraturlesnad*. 6. utg. Oslo.
- Utne, Ivar. 1991a. Kva kan ein vente av eit program for stavekontroll? I: *Språknytt* nr. 2 1991.
- Utne, Ivar. 1991b. Val av språknorm i dataordlister. I: *Språknytt* nr. 3 1991.
- Warren, Dagfinn: 1988. Berre ord. I: *I klartekst. Festskrift til Bernt Fossestøl på 60-årsdagen*. Oslo.

Jørn-Otto Akø

## Gråsoner i norske ordbøker

Under arbeidet med ei maskinleselig ordliste for bokmål blei det oppdaga en del uklarheter, notasjonsfeil og systematiske mangler i de to nyeste ordbøkene, *Tanums store rettskrivningsordbok* og *Bokmålsordboka*. De punktene som burde ha generell interesse både for andre ordboksredaktører og for vanlige brukere, blir gjennomgått, og det blir skissert enkelte mulige løsninger.

### 1 Innledning

Bakgrunnen for denne artikkelen er arbeidet med å redigere den bokmålsordlista som ligger til grunn for blant annet stavekontrollen i IBMs tekstbehandlingsprogram.

#### 1.1 Stavekontroll

I tradisjonelle ordbøker som blir gitt ut som papirutgaver, behøver ikke redaktørene presisere alle mulige bøyingsopplysninger. De kan la brukeren bestemme om substantiv kan brukes i flertall, om adjektiv kan gradbøyes osv., og de kan gi generelle regler og retningslinjer i forordet.

En stavekontroll fungerer derimot bare på grunnlag av de opplysningene som ligger til grunn for den, og de må være riktige og entydige. Det betyr at vi som redigerer ordlistene som ligger til grunn for stavekontrollen, må bestemme om et substantiv bare skal ha entallsformer, om det bare skal ha flertallsformer, om det skal ha begge deler, eller om det er ubøyelig, om et verb kan ha attributive former (se nedafor) og hvordan disse i tilfelle skal se ut, om adjektiv kan gradbøyes med suffiksene *-ere/-est* osv.

Hensikten med stavekontrollen er at brukeren fortløpende skal få kontrollert om de ordformene hun bruker er korrekte, uten å måtte forlate PCen eller terminalen for å slå opp i ei ordbok.

Når vi lager en stavekontroll, kan vi velge å ta med bare de ordformene som vi ut fra frekvensstudier veit er de mest brukte. Problemet med denne strategien er at brukeren vil undre seg over at noen ordformer av et leksem blir akseptert som riktige former, mens andre former, som også er riktige, *ikke* blir akseptert.

En annen løsning er å ta med hele paradigmet når vi først tar med ei ordform. Da vil vi altså ha med *alle* riktige bøyingsformer av *hvert* oppslagsord.

## 1.2 Utvalg av oppslagsord og bøyingsopplysninger

Når vi lager ei maskinlesegelig ordliste, har vi det samme problemet som redaktører av andre ordlister og -bøker når det gjelder utvelgning av hvilke oppslagsord som skal være med. Fordi det hele tida kommer nye ord og uttrykk inn i språket, og fordi det ikke er praktisk mulig å ekserpere ord fra alle (fag)felter, kan vi aldri klare å få med alle de orda brukerne bruker.

Normen er til enhver tid nedfelt i de nyeste ordbøkene, og de to viktigste for bokmål er *Tanums store rettskrivningsordbok* (Tanum) og *Bokmålsordboka* (BOB). Vi som lager stavekontroller, er også henvist til å slå opp i dem for å kontrollere at skrivemåter og bøyingsendelser er riktige.

Finner vi så det vi trenger i ordbøkene? Som oftest gjør vi sjølsagt det, men når vi ikke finner de opplysningene vi trenger, er det all grunn til å vente at de fleste ordboksbrukerne får det samme problemet. Av og til fører Tanum og BOB opp ulik rettskrivning eller ulike bøyingsformer uten at forskjellen skyldes at det har vært normendringer i mellomtida. Hos IBM har vi i flere år ført en fortløpende oversikt over alle uoverensstemmelser mellom de to ordbøkene, og fordi vi har så gode hjelpemidler til å sortere og gruppere ord og bokstavstrenger, finner vi mange uoverensstemmelser som ellers neppe ville ha blitt oppdaga. Norsk språkråd normerer på grunnlag av de orda de har oversikt over, men fordi man heller ikke der har tilstrekkelig kapasitet til å hente ut komplette lister over aktuelle ord, oppstår det av og til enten "hull" i listene, eller det blir forskjellig normering av ord som naturlig hører sammen. Eksempler på dette er *brasilianer*, som ikke hadde kommet med i lista over normerte nasjonalitetsbetegnelser i Årsmeldinga for 1983, og *avgift*, som bøyes med *-a* el. *-en*, mens *utgift* bøyes med *-en*. Vi håper at slike oppdagelser kan bidra til enkelte forandringer og utvidelser i seinere utgaver av bokmålsordbøkene.

Opplysninger om gradbøyning av adjektiv og adverb, attributive former av verb, numerus av substantiv osv. blir i mange tilfeller overlatt til språkbrukerens egen språkfølelse. Problemet med denne følelsen er at den ofte vakler, at den varierer fra person til person, og at personer som ikke har norsk som morsmål, ikke har noen språkfølelse å støtte seg til. Disse gruppene har altså felles interesse av at så mange opplysninger som mulig blir skrevet fullt ut. Det ideelle ville være at alle relevante opplysninger stod i hver ordboksartikkel, men av reine plasshensyn går jo ikke det.

Når vi først har valgt oppslagsordet, har vi med en endelig mengde bøyingsopplysninger å gjøre. Da gjenstår "bare" problemene med hvilken notasjon man skal velge, og hvilke bøyingsopplysninger som skal tas med for hver enkelt ordklasse.

Videre kan vi skille mellom to ulike strategier for å *føre opp* bøyingsopplysninger. Når det gjelder attributive former av verb, kan man velge å ikke angi de attributive formene under hvert oppslagsord, men formulere regler i forordet som viser hvordan formene skal se ut. Dette er den løsningen redaksjonen av Tanum har valgt, men når det gjelder "Bøyning av adjektiv og partisipp på -et" som de kaller det i forordet, blir reglene så vage at en vanlig bruker iallfall ikke klarer å bruke dem riktig.

En annen løsning ville være å spesifisere de attributive formene under hvert enkelt oppslagsord. Da vil fraværet av denne bøyingsopplysningen bety at forma ikke fins. Dette er en god løsning for brukeren fordi hun da har "alle" opplysningene tilgjengelig for hvert enkelt oppslagsord. For redaktørene betyr det sjølsagt betydelig merarbeid, men her blir det snakk om hvor mye hjelp vi ønsker å gi brukeren.

## 2 Gråsoner

Jeg velger her å konsentrere meg om attributive former av sterke verb (2.1) og bare kort skissere andre problemområder.

### 2.1 Attributive former av verb

Attributive former av verb, altså opphavlige partisipper som blir brukt som adjektiv, er et av de største problema vi har i bokmålsmorfologien. Grunnen til at folk har problemer med disse formene, er at spriket mellom normen i skriftspråket og det talemålet folk flest bruker, her er større enn på de fleste andre områder. Når folk lurer på om de skal skrive *brukkede* eller *brukne*, er det fordi de sier *brekte* eller tilsvarende.

Vi finner også verb der bare den ene perfektum partisippen kan brukes som attributiv form:

løpe, løp, løp(e)t

en løpt distanse

\*en løpet distanse

Det hersker en utbredt forestilling om at transitive verb *har* attributive former, mens intransitive verb *ikke* kan ha det, men slik er det ikke. Den eneste gruppa med verb som bare sjelden kan ha attributive former, er reint refleksive verb som for eksempel *skamme seg*, og verb som *forstrekke*, som kan være både transitive (*forstrekke en muskel*) og refleksive (*forstrekke seg*). Vi finner unntak som *forelske seg* og *forspise seg* som kan ha attributive former:

en forelska gutt

en forspist middagsgjest

*Rekke* er et godt eksempel på transitivt verb som *ikke* kan danne attributive former:

rekk meg en svamp

\*en rukkit svamp

*Utebli* og *slokne* er gode eksempler på intransitive verb som *kan* danne attributive former:

en uteblitt deltaker

en slokna el. sloknet brann

Det er altså i hvert fall ikke transitivitet aleine som avgjør hvilke verb som *kan* danne attributive former.

Spørsmålet blir da om det ikke burde markeres i ordbøkene hvilke verb som har attributive former, og hvilke som ikke har det. Hvis vi innfører attributive former som en obligatorisk kategori ved oppføring av verb i ordbøkene, får vi imidlertid problemer med homonymi/polysemi. Hvordan skal vi føre opp verb som kan ha attributive former i bare en

eller noen betydninger, og skal verb som kan forekomme både med og uten attributive former føres opp som to oppslagsord?

### 2.1.1 Perfektum partisipp

De største "gråsonene" finner vi ved perfektum partisipp. Innafor denne gruppa igjen er det hankjønns- og hunkjønnsformene av de sterke verba som er mest problematiske. I bokmål har vi nå tre grupper med opprinnelige perfektum partisipp og analoge dannelser som utgjør en glidende overgang fra reine partisipp til reine adjektiv<sup>1</sup>.

#### a) Reint partisipp:

skrive, skrev el. skreiv, skrevet

mask/fem:	en skrevet henvendelse
nøytrum:	et skrevet testamente
flertall/svak form:	skrevne brev/ den skrevne henvendelsen

#### b) Blanding av partisipp og adjektiv:

binde, bandt, bundet

mask/fem:	bunden el. bundet
nøytrum:	bundet
flertall/svak form:	bundne

*bunden el. bundet* og *grepen el. grepet* er blant de ytterst få tilfellene i ordbøkene der de attributive formene av et usammensatt sterkt verb har egen oppføring.

#### c) Reint adjektiv:

mask/fem:	frossen
nøytrum:	frossent
flertall/svak form:	frosne

(verbet som ligger til grunn her, er *fryse, fraus el. frøys, frosset*)

#### 2.1.1.1 Usammensatte verb

Usammensatte verb blir ikke behandla helt på samme måte som de sammensatte i ordbøkene.

Reine adjektiv av opprinnelig usammensatte sterke verb er greie nok, sjøl om de ikke viser om verbet eventuelt har attributive former i tillegg. Når Tanum fører opp adjektivet

---

<sup>1</sup> Årsmelding fra Norsk språkråd 1983, side 21.



*sprukken* i tillegg til partisippet *sprukket*, så er det ikke lett for en vanlig bruker å skjønne om både adjektivet og partisippet kan brukes attributivt.

Andre partisipper av sterke verb blir altså bare ført opp som perfektum partisipp, og her er ikke den gjengse språkfølelsen helt i takt med normeringa. Mange brukere er usikre på hvilke attributive former vi kan ha av partisipper som:

sprukket  
brukket

Her vil nok mange brukere vakle mellom følgende former: .

sprukken; sprukken el. sprukket; sprukkede  
brukken; brukken el. brukket; brukkede

Lettere blir det ikke for brukerne hvis de oppdager at samme verb blir ført opp forskjellig i forskjellige ordbøker. I tillegg til verbet *sprekke* finner vi følgende oppføringer:

- 1) Tanum: sprukken, adj., n. -t; sprukne
- 2) BOB: sprukken adj sprukket, sprukne (av sprekke)

Her vil en vanlig bruker måtte trekke den slutningen at de to ordbøkene er uenige om hvordan ordet skal bøyes.

- 3) Vinje: *Moderne norsk* (1987, s. 120, spalte 2 og s. 121, spalte 2):

sprukken el. sprukket, sprukket, sprukne (egentlig, opprinnelig betydning)  
sprukken, n. -t, sprukne (uegentlig, overført betydning)

Som eksempler fører Vinje opp *et sprukket klemt* - *et sprukket glass*. Han skiller altså mellom to ulike betydninger. Tanum har med den ene, mens BOB fører opp ei form i maskulinum og femininum som ikke var i samsvar med Språkrådets normering på det tidspunktet BOB blei gitt ut.

De attributive formene av verbet *briste* lar seg ikke utlede av ordbøkene. Verbet bøyes: *briste, brast, bristet* el. *brustet*. Ifølge Språkrådet skal *briste* ha følgende attributive former:

mask/fem:	bristet
nøytrum:	bristet
flertall/svak form:	bristete, bristede

Eks: en bristet muskel, et bristet ribbein, bristete el. bristede ribbein

I tillegg har vi et reint adjektiv *brusten* som bøyes slik:

mask/fem:	brusten
nøytrum:	brustent
flertall/svak form:	brustne

Eks: en brusten forhåpning, et brustent hjerte, brustne forhåpninger

Ut fra systemet i bokmåls morfologien kunne vi også ha tenkt oss at maskulinum og femininum hadde hatt formene *brustet*, eller *brusten el. brustet*.

Usikkerheten henger her kanskje delvis sammen med ei normeringsendring som Språkrådet foretok i 1983. Da blei de stivna partisippene redusert fra fem til tre grupper. Dessverre har Språkrådet bare tatt med noen få eksempler i Årsmeldinga for dette året, og mange av de orda som kommer inn under dette punktet, står oppført med feil bøyning i både Tanum og BOB. Et godt eksempel er *bred- el. breiskåren* som i Tanum blir ført opp uten spesifisert bøyning. Da skulle vi vente at disse orda skulle bøyes som perfektum partisipp av det usammensatte verbet *skjære*. Partisippet har egen oppføring som attributiv form:

skåren el. skåret, adj., n. ..et; fl. ..rne,

men ifølge referatet fra et møte i fagnemnda i Språkrådet 20.6.1983 skal *bred- el. breiskåren* bøyes:

bred- el. breiskåren, n. -skårent; fl. -skårne

### 2.1.1.2 Avleda og sammensatte verb

Når det gjelder attributive former av avleda og sammensatte sterke verb, fører BOB opp alle de riktige bøyingsformene for de oppslagsorda som har blitt tatt med, men fordi mange verb ikke blir ført opp der, må vi ty til Tanum, som har flere oppslagsord. Tanum fører som regel bare opp ei attributiv form i maskulinum og femininum, for eksempel

innknepet, oppknepet,

men problemet er at det i mange tilfeller fins flere riktige former enn den ene som står oppført. Vi kan altså ikke stole på at den forma som står i Tanum alltid er den eneste riktige.

I tillegg har Tanum oppføringer som ikke stemmer overens med vedtak i Språkrådet. Følgende oppføringer i Tanum er tilsynelatende like:

innknepet, adj., fl. ..pne  
oppknepet, fl. ..pne adj ("om hest med sterkt opptrukket buk")

Den mest nærliggende tolkinga av disse oppføringene er:

mask/fem:	innknepet; oppknepet
nøytrum:	innknepet; oppknepet
flertall/svak form:	innknepne; oppknepne,

men Språkrådet har normert disse orda slik i maskulinum og femininum:

mask/fem:	innknepen el. -et
	oppknepet

For en vanlig bruker er det umulig å gjette seg til at like oppføringer ikke alltid skal leses på samme måte.

Videre er det vanskelig å forstå hvorfor Språkrådet mener at det i maskulinum og femininum skal hete *oppsprukken* el. *oppsprukket*, men derimot bare *opptrukket*. Her har vi samme type partisippdannelse, og det er verb av samme bøyingsklasse som ligger til grunn.

## 2.1.2 Presens partisipp

Den andre gruppa med attributive former er presens partisipp. Denne forma danner vi av infinitiv pluss endelsen *-(e)nde*. De to ordbøkene fører ikke konsekvent opp presens partisipp, verken som egne oppslagsord eller som bøyingsopplysninger under verbet, mens Tanum fører opp ing-former som en av bøyingsformene av verbet.

Presens partisipp blir normalt danna på samme måte som ing-formene, men det står det ingenting om i ordbøkene, heller ikke i forordet. Problemer på dette området får vi først når vi vanskelig kan se hva som er verbets stamme, eller når presens partisipp og ing-former blir ført opp på ulik måte.

### 2.1.2.1 Verb med kortformer

Det første problemområdet her er verb som har, eller har hatt, både ei kort og ei lang form, for eksempel:

ri el. ride  
fly el. flyge

Tanum fører opp slike verb på mange ulike måter:

a: li el. lide, ..., -ding  
b: lide el. li, ..., -ing  
c: bestri(de), ..., -ing  
d: kle el. klede, ..., -(d)ing  
e: miskle, ..., -(d)ing  
f: ska, ..., -ding, el. skade, ..., -ing  
g: gjø, ..., -ing ("om hund")  
   gjø(de), ..., -ing ("feite")  
h: nø, ..., -ing el. nøde, ..., -ing

Følgende notasjoner er ikke entydige:

b: lide el. li, ..., -ing: gir *liding* av *lide* og *liing* av *li*  
d: kle el. klede, ..., -(d)ing: gir *kleing* el. *kleding* av *kle* og *kleding* av *klede*

Disse notasjonene er etter min mening entydige:

a: li el. lide, ..., -ding : gir *liding* av begge verba

c: bestri(de), ..., -ing: gir *bestriing* av *bestri* og *bestriding* av *bestride*  
 e: miskle, ..., -(d)ing: gir valgfritt *miskleing* og *miskledning* av *miskle*  
 f: ska, ..., -ding, el. skade, ..., -ing: gir *skading* av både *ska* og *skade*  
 g: gjø, ... -ing ("om hund"): gir *gjøing*  
     gjø(de), ... -ing ("feite"): gir *gjøing* av *gjø* og *gjøding* av *gjøde*  
 h: nø, ... -ing el. nøde, ... -ing: gir *nøing* av *nø* og *nøding* av *nøde*

Verbet *li* får altså ulik ing-form under de to oppføringene. Her er det vanskelig for en vanlig bruker å se hvilken ing-form det enkelte verbet skal ha, og følgelig også hvordan presens partisipp av det samme verbet skal være. Enda et problem har vi når det gjelder presens av slike kortverb. Enkelte mener at ordforma *rider* kan være presens av både *ri* og av *ride*, og det er et godt argument for å ta presensformene av disse verba med i ordbøkene, slik at brukeren ikke er i tvil om hvilke former som er riktige.

### 2.1.2.2 Presens partisipp er ulik ing-forma

Vi finner også tilfeller der presens partisipp og ing-forma blir ført opp på forskjellige måter:

gi, ..., -ving, men: gi(v)ende

I ing-forma skal vi altså ha med den opprinnelige rotkonsonanten, mens vi i presens partisipp kan velge om vi vil ha den med. Uansett om slike forskjeller er motiverte eller ikke, vil ordboksredaktørene bli tvunget til å ta stilling til dem hvis presens partisipp og ing-form blir obligatoriske deler av artiklene. Med en slik løsning vil også manglende oppføring av presens partisipp og ing-form bety at det aktuelle verbet ikke kan ha den ene eller den andre forma.

### 2.1.2.3 Verb med "flytende j"

Verb med "flytende j" er et spesielt problem i forbindelse med oppføringene i begge ordbøkene. Det er ikke mange av dem, men de skaper mange problemer fordi oppføringene ikke inneholder alle de bøyingsformene brukerne er usikre på. I Tanum finner vi disse gruppene:

a: verb som mister j-en i preteritum:

følgje, følgte, følgt; -ing  
 røkje, røkte, røkt; røking ("undersøke")  
 svelgje, svelgde, svelgd; svelging  
 telgje, telgde, telgd; telging

*Følgje* gir altså *følgjing*, mens de andre verba danner ing-form uten j-en.

b: verb som beholder j-en i preteritum:

bølge, ... el. bølge, -a/-et, -a/-et; -ing

*bølge gir bølging, mens bølge gir bølging.*

Hvis oppføringene i Tanum er riktige, bør redaktørene for det første skrive ing-formene fullt ut for disse verba. I tillegg burde de ta med presens og presens partisipp for å vise hvilke former som er riktige. Hvis oppføringene i Tanum er feil, må redaktørene rydde opp i samarbeid med Språkrådet.

### 2.1.3 Oppsummering

- Ing-formene bør tas med som ei bøyingsform av verbet også i BOB.
- Ing-former av verb med kortformer og verb med "flytende j" bør skrives fullt ut slik at oppføringene ikke kan misforstås.
- Presens partisipp bør tas med som ei av bøyingsformene av verbet. En mulighet er å ta med presens partisipp ved hvert oppslagsord, og la fravær av presens partisipp bety at forma ikke bør brukes. En annen mulighet er å nevne i forordet at presens partisipp dannes på samme måten som ing-formene, og markere de verba som sjelden eller aldri kan danne presens partisipp.

## 2.2 Numerus av substantiv

Tanum og BOB bruker ulike strategier når de fører opp numerus av substantiv. Tanum fører i mange tilfeller bare opp entallsforma, sjøl om de fleste ville være enige i at ordet kan brukes i flertall også, jamfør lista nedafør. Dette medfører særlig problemer ved nøytrumsord og ord som kan være både maskulinum og nøytrum. Her er det vanskelig for en vanlig bruker å vite om ubestemt flertall skal være -, - el. -er eller -er:

	Tanum	BOB
anke	-n	m1
beta	-en	m1
buké	-en	m1
bouquet	-en	m1
delta	-en	m1 ("gresk bokstav")
giss	-en	m1
miks	-en el. -et	m1 el. n1
støp	-en el. -et	m1 el. n1
sortiment	-et	n3
øl	-et	n1 ("ølsort")

BOB fører svært sjelden opp bare entallsforma, men når den gjør det, er det som regel motivert; det vil si at det er vanskelig å tenke seg flertallsformer av oppslagsordet. Annerledes er det når det blir brukt vanlig kode. Kodene m1, f1 osv. sier ikke alltid at oppslagsordet kan brukes i flertall, men de sier i det minste hvilke bøyingsendelser orda skal ha i flertall hvis de kan brukes.

Når Tanum fører opp bare entallsformer av et oppslagsord, har ikke det en enhetlig betydning. Noen av disse oppslagsorda blir brukt bare i entall, noen kan opplagt brukes i

flertall, noen kan kanskje brukes i flertall, men hvordan skal formene i tilfelle se ut? Her burde redaktørene prøve å få til en enhetlig bruk av notasjonen.

### 2.3 Trekjønna substantiv

Enkelte substantiv kan være av alle tre kjønn, men måten disse orda blir ført opp på i Tanum, gjør det vanskelig for en vanlig bruker å se hvor mange oppslagsord det er snakk om:

lære, ..ra el. -n; -r, -ne el. lære, -t; -r, ..ra el. -ne ("mal; måleverktøy")  
 spenne, ..na el. -n; -r, -ne el. spenne, -t; -r, ..nna el. -ne  
 nyre, ..ra el. -n el. -t; -r, ..ra el. -ne  
 dun, -a el. -en el. -et; -, -a el. -ene

At *dun* og *nyre* er trekjønna ord, vil nok de fleste være enige i, men det er klart at vi ved *lære* og *spenne* har ett oppslagsord som er maskulinum og femininum og ett som bare er nøytrum. Her bør redaktørene velge en enhetlig og entydig notasjon.

### 2.4 Bestemt form flertall mangler

Dette gjelder særlig engelske lånord som *gallon*, *mile*, *royalty* og måleenheter som *ohm*, *ampere*, *watt*, *volt*. Bestemt flertall er formodentlig utelatt fordi redaktørene har ment at formene er regelmessige, og at de kan utledes av de andre formene. Slik ser oppføringene ut i dag:

	Tanum:	BOB:
gallon:	en; -el. -s <sup>2</sup>	-en, - el. -s
mile:	en <sup>3</sup> ; -s	(ikke oppført)
royalty:	-en; fl. ..ties	-en, -ties
ohm:	-en; -	m l <sup>4</sup>
ampere:	-n; fl. -	-n, -
watt:	-en; fl. -	-en, -
volt:	-en; -, -ene	-en, -

Da vi tok opp disse orda med Språkrådet, fikk vi følgende normeringer:

gallon:	-en; fl. - el. -s, -ene
mile:	-n; fl. -s, -ne
royalty:	-en; fl. royalties el. royaltyer, royaltiene
ohm:	-en ; fl. -, -ene
watt:	-en; fl. -, -ene

<sup>2</sup> Uten prentesvartens herjinger ville nok oppføringa ha sett slik ut: gallon, -en; - el. -s.

<sup>3</sup> Her skulle det trolig stått -en.

<sup>4</sup> Koden m l gir formene -en; -er, -ene.

Som vi ser, følger ikke alle formene i bestemt flertall av de andre formene som er ført opp i ordbøkene. Ved *royalty* ser vi at ubestemt flertall ikke er helt enkelt heller. For *ohm* fører BOB opp en bøyingskode som gir feil form i ubestemt flertall, men det at forma *er* angitt, gjør det mulig å vurdere om den er riktig. Jeg mener at redaktørene burde ta med *alle* bøyingsformene fordi:

- brukeren da slipper å utlede bøyingsformene sjøl
- det at bøyingsformene faktisk *står* i ordboka, er en forutsetning for at brukerne skal kunne vurdere om de er riktige.

## 2.5 Gradbøying

De fleste er vel enige i at et adjektiv som *vovet* må gradbøyes med *mer* og *mest*, men slike opplysninger finner vi ikke i ordbøkene. Det er bare ved adjektiv som får suppletivistiske former i komparativ og superlativ, som for eksempel *liten*, *mindre*, *minst*, at vi finner opplysninger om gradbøying. De andre adjektiva blir overlatt til den enkeltes skjønn med de konsekvenser det får. Hvilke av de følgende adjektiva som kan gradbøyes, overlater jeg til leserne:

abstrakt  
 absolutt  
 likvid  
 naivistisk  
 promiskuøs  
 statsfiendtlig  
 søkt ("underlig, snodig; spesiell")  
 utilpass  
 vidgjeten

Jeg mener at redaktørene burde ta med opplysninger om gradbøying. Det blir en stor og vanskelig jobb, men det ville gi brukerne en bra veiledning, og det ville gi de spesielt interesserte muligheten til å diskutere om de formene redaktørene har kommet fram til, er de riktige.

## Litteratur

- Fossestøl, Bengt og Sandvei, Marius (red). 1989. *Tanums store rettskrivningsordbok*.  
 Landrø, Marit Ingebjørg og Wangensteen, Boye (red). 1986. *Bokmålsordboka*. Oslo.  
 Vinje, Finn-Erik. 1987. *Moderne norsk*. Oslo.

## Om leksikografi som undervisningsemne

Leksikografi er et fag i framgang. Det har hittil vært lite av systematisk undervisning i faget; tiden bør nå være inne til å lage faste undervisningsopplegg på universitetsnivå og som del av embetseksamen. Et halvårig kurs ("semesteremne") har vært tilbudt ved Universitetet i Oslo siden 1986 og kom i gang fra 1988. På grunnlag av de tre første kursene beskriver artikkelen opplegget, problemer ved faget og ved pensumlitteraturen av særlig betydning for undervisningen, og studenters vurdering av kurset.

### Innledende ord om hvorfor

De nordiske land har gjennom flere hundre år fostret en lang rekke fremragende leksikografer. Så vidt meg bekjent hadde ingen av dem formell utdanning i et fag som ble kalt leksikografi, enda ordet skal være belagt i fransk fra 1578, i engelsk fra 1680, i tysk fra 1698 (*Lexikograf* fra 1641) og i svensk (også her *lexikograf*) fra 1780.

Dette er en tanke som kan mane til ydmykhet overfor det å skulle etablere leksikografi som undervisningsfag. Blir man en bedre leksikograf av å studere leksikografi? Ville de store pionerene blitt større om de hadde gjort det?

Det er ikke sikkert. Derimot tør man tro at hvis de ved begynnelsen av sin leksikografiske karriere kunne studert fagets teori og praksis, ville de blitt selvstendige leksikografer *raskere* enn de ble det.

Til daglig arbeider leksikografer mest med materiale som ikke er helt nytt ("nyord"), men som har vært under behandling av andre, i det minste skaperen av den primærkilden vi har hentet ordet fra. Råstoffet er delvis bearbeidet, iallfall ortografisk, og egentlig også semantisk gjennom forfatterens valg. Dette kan tyde på at man i vårt fag i sterkere grad enn i mange andre står på skuldrene til sine forgjengere; det er et fag der det å ta stilling til hva disse har gjort, inngår mer aktivt i arbeidsprosessen enn bare å beskrive "Stand der Forschung" eller "state of the art", og at faget derfor egner seg for å læres ved at man som disippel blir innført skritt for skritt av noen som har gått veien før.

Men det er som kjent en lang vei. Det finnes sitater, vers om den unge og den gamle hr. Brandt og bevingede ord av Samuel Johnson og andre som fastslår dette. Det finnes en tabell i Zgustas *Manual of Lexicography* (1971:349) over antall år og prosent som planen for utgivelsen av fire store europeiske ordboksverk ble overskredet med: 360 % for det som kom best ut, 640 % for det som kom dårligst ut til da; det er ikke ferdig ennå. Det er nærmest et aksiom i vårt fag at jo mer man vet om ordboksarbeid, jo mer pessimistisk ser man på muligheten til å overholde en planlagt ramme for utgivelsen, hvis man da ikke først har beregnet antatt tid og så lagt inn et passende antall år ekstra, etter hva erfaring tilsier, før man gjorde beregningen kjent.



At det er slik, kan bringe pessimismen frem. Lar ordboksberegning seg overhodet lære, når en vet at alle anser det som viktig, prøver å greie det, men likevel mislykkes? En kunne trenge å være utdannet forlagsredaktør, typograf og økonom i tillegg til sitt egentlige fag, og dessuten kjenne noen grufulle hemmeligheter i livet til visse personer i bevilgningsposisjon, så en kunne presse de nødvendige bevilgninger ut av de motvillige.

Tross alle ufullkommenheter er det likevel mye som taler for å etablere en systematisk undervisning i leksikografi. For å si det bibelsk, så er høsten stor, men arbeiderne få. Det skulle man kanskje ikke tro her og nå, men vi kan modifisere det til at arbeidstimene er få, for det er ikke så mange av oss som har anledning til å drive faget på full tid.

For øyeblikket har vi i Oslo ca. 75 fremmedspråk representert i grunnskolen, med ganske små brukergrupper - ned til én. I vårt lille språksamfunn har det hittil vært ytterst vanskelig å utgi fremmedspråklige ordbøker for små grupper. Økonomien er vel den største hindringen, men også at ordboksinitiativet ikke sjelden blir tatt av en enkeltperson som ser behovet, begynner hjemme på kjøkkenet, og har et totalt ubehjelpelig forhold til ordboksredigering. For øvrig er det eiendommelig å se hvor ofte også godt skolerte folk, tilmed lingvister, som setter i gang ordboksarbeid uten å ha lært det, gjør alle de elementære redigeringsfeil som det er mulig å gjøre. Hva disse går ut på, er det unødvendig å regne opp her, men la meg likevel nevne dem, som et raskt gjensyn med gamle venner:

- 1 sirkeldefinisjon ( $a = b$ ,  $b = a$ )
- 2 ulik ordklasse i oppslagsordet og hovedordet i definisjonen
- 3 for snever eller (sjeldnere:) for vid definisjon
- 4 definisjonen er vanskeligere enn oppslagsordet selv
- 5 oppslagsordet, eller en avledning av det, inngår i definisjonen
- 6 nøkkelord i definisjonen mangler som egne oppslagsord
- 7 oppslagsordet og definisjonen har ulik bestemthet eller ulikt tall
- 8 vakling mellom substitusjonsprinsippet og encyklopediprinsippet, dvs. mellom en definisjon som kan erstatte oppslagsordet, og en forklaring i full(e) setning(er)<sup>1</sup>

Vårt eget språk er i raskere endring enn før. Ikke bare er vi nordboere store importører av fremmede ord. Vi skifter også ut viktige kilder til vår egen språkbruk. Bibelen, nasjonallitteraturen og primærmæringerne er på retur som språklig arsenal. Det tradisjonelle fondet av kollokasjoner, idiomer og andre språklige bilder blir oftere feilbrukt i form, betydning og sammenheng. Det trengs at noen holder en finger på språkets puls, blant annet for at presisjonsnivået ikke skal synke enda mer. Det får vi hyppige eksempler på i avisspråket.

Her har vi et problem: Leksikografen deltar i normeringen av språket og fastslår hva som er riktig. Han søker da en rettesnor ved å se bakover - noe er riktig i dag fordi det var riktig i går. På den annen side registrerer han faktisk språkbruk, der bruddene på de tradisjonelle regler florerer, og der nye ord, betydninger og bruksmåter dukker opp. Hvordan skal han forholde seg til dette? Hvis registreringen skal brukes til ny normering, inngår det å spå om fremtiden: Hva er riktig i morgen? Hvor går trenden? Skal vi f.eks. begynne å akseptere løs sammensetning etter engelsk mønster, *mann-maskin dialog* og *langdistanse radar*, fordi det er blitt så vanlig, og fordi fast sammensetning ofte gir en uhandterlig ordlengde, eller i motsatt fall uforståelige initialord? Er det mulig å lære studentene noe om feil som er og foreløpig blir uakseptable, og feil som vil inngå i neste tiårs språkvaner og kommer inn i ordbøkene? Det er trekk ved vår språksituasjon og -utvikling som minner om

Heisenbergs usikkerhetsrelasjon anvendt på legemer i bevegelse: Man kan ikke samtidig fastslå posisjon og hastighet.

### Et opplegg

Etter å ha brakt på bane disse store og til dels eksistensielle spørsmål, er jeg klar til å presentere et undervisningsopplegg for begynnerstudenter i leksikografi på 12 uker à 3 timer pluss oppgaveskriving og en avsluttende muntlig prøve. Det er en begynnelse. Forhåpentlig kan det utbygges etter hvert. Inntil så skjer, må studiet gjøres så nyttig som mulig innenfor den rammen vi har.

Jeg nevner dette for å gjøre det klart at vårt opplegg er meget beskjedent. Lenger er vi ikke kommet. Jeg håper at interesserte som eventuelt ikke engang er kommet så langt, kan få noen ideer til hvordan de vil gjøre det, eller nettopp ikke gjøre det, eller at de i det minste kan legge bort all skepsis til å begynne i det små.

Ved årsskiftet 1984/85 nedsatte vårt Fakultet en komité til å utarbeide forslag til studieplan for et semesteremne i leksikografi, altså en halvårsenhet. Det samme ble gjort med flere andre "små" fag for å trekke dem inn i undervisning og formidling. Vårt forslag var ferdig neste vår, det ble godkjent etter enda et år, og emnet ble tilbudt studentene fra høsten 1986; men det første kurset kom i gang først i 1988. Vi tilbyr det bare hvert høstsemester. Hittil har vi holdt tre kurs, gjennomført av 14 personer, som altså, i motsetning til sine egne forelesere, nå har en formell utdanning i faget og eksamenspapir på dette.

Som forelesere har flest mulig av vårt eget personale vært trukket inn. Det skaper bedre intern oppslutning om formidlingssiden ved faget. At man som ordboksarbeider blir ansporet til å formidle det man driver med, formulere det for seg selv, og ikke minst knytte det til den bakenforliggende teorien, gir også opplevelser som ikke er å forakte. - Videre trekker vi i noen grad inn eksterne krefter; bl.a. gis det en kort innføring i terminologi, likeså i emner fra et fremmedspråk som noen av studentene arbeider med eller foretrekker - hittil har rekkefølgen vært spansk, engelsk og tysk.

Siden studiet skal svare til ett semesters fullt arbeid, er pensum på ca. 1350 sider. Det er delt i tre avdelinger: allmenn leksikografisk teori, ettspråklig, særlig norsk og nordisk leksikografi, og tospråklig (fremmedspråklig) leksikografi. Hovedvekten ligger på norsk, deretter på nordisk og så på fremmedspråk. Det inngår ikke i vårt arbeidsfelt å behandle fremmedspråk uten tilknytning til norsk enten som opphavsspråk eller målspråk; vi lager f.eks. ikke ordbøker fra ett fremmedspråk til et annet.

Før studentene tar vårt kurs, kreves det av dem at de har tatt en fageksamen på universitetsnivå enten i nordisk eller et fremmedspråk. Fra dette kravet kan det gjøres unntak, f.eks. hvis en edb-kyndig student fra et annet miljø enn vårt ønsker å arbeide med den leksikalske siden av et fagspråk. Dette har hittil ikke vært aktuelt i praksis; derimot har flere av dem som fylte opptakskravene, tatt kurset av interesse for fagspråk. Dette er da gjerne modne studenter som bidrar til en verdifull dialog.

I løpet av kurset skal studenten besvare tre kortere skriftlige oppgaver og utarbeide en noe lengre semesterrapport. De korte oppgavene kan f.eks. være å vurdere en trykt ordbok, ekserpere eller kode relevant ordboksmateriale fra utlevert primærmateriale, eller skrive holdbare definisjoner på grunnlag av bilder eller av ord i kontekst i primærkilder. Rammen for semesterrapporten trekkes meget vidt. Blant innleverte arbeider hittil kan jeg nevne:

Norske avløserord for anglismer  
 Ordavledninger på *-eri*, særlig brukt nedsettende  
 Ord i norsk med *fritid* som forledd  
 Bruk av ordbøker ved oversettelse fra spansk  
 Norske militære ord og uttrykk med franske ekvivalenter  
 Adjektiviske fargebetegnelser i Henrik Wergelands skrifter  
 (Flere har beskrevet lydverk og ordforråd i en dialekt)

Emnevalget reflekterer undervisningen og det materialet avdelingen kan tilby, men kanskje særlig studentenes egne interesser, idet vi oppmuntrer dem til å finne et brukbart emne selv, før vi eventuelt gir dem en oppgave og materiale til å løse den. Blant oppgitte emner skal jeg siden komme tilbake til ett: Høsten 90 skrev en student oppgaven "En evaluering av semesteremnet i leksikografi"; vi bad ham gjøre det etter at han hadde vist interesse for å kommentere kursopplegg og pensum.

Kurset avsluttes med en kort muntlig prøve. Den "skriftlige" eksamen tas altså underveis i form av oppgaveskriving. Sluttresultatet er etter Fakultetets pålegg enten "Godkjent" eller "Ikke godkjent" - det siste har ikke forekommet.

### **Pensum, innhold**

Innenfor hver av de tre avdelingene av pensum er en del obligatorisk og en del valgfri; det obligatoriske stoffet utgjør litt over halvparten. Vi har satset mye på at all nødvendig litteratur skal være å finne i vårt eget bibliotek, og bibliotekets hyllemarkering er innført i pensumlisten ved hvert verk. I tillegg deler vi ut en ukeplan over undervisningsemne(r) for hver uke, med en oppstilling av pensumlitteratur som passer til ukens emne(r). Dette gjør det lett å finne frem i stoffet, noe studentene er fornøyde med. Slik pensum ser ut, er det også ganske nødvendig.

Kurset starter med en innføring i semesteremnet, pensum, opplegg, vår avdeling, dens oppgaver, materiale og hva vi kan tilby. Av særlig interesse er det å gå litt i dybden av vår nyordsinnsamling. "Nyord" vil si nyregistrerte ord, betydninger og bruksmåter. At selve ordet virkelig er nytt, vet en jo sjelden sikkert.

For å gjøre materialet så allsidig brukbart som mulig, koder vi det grammatisk, eventuelt syntaktisk og pragmatisk eller typologisk alt etter ordets karakteristiske trekk. Vår kodeliste med kommentar kan ses som en ilmarsj gjennom språket, eller et "språkbad", som noen språkpedagoger har begynt å kalle det.

I den andre enden av kurset har vi en innspurt der kursinnholdet blir repetert på langs; et siste forsøk blir gjort på å besvare spørsmål og oppklare dunkle punkter og hull i studentenes viten. For å identifisere hull kan de selv bli bedt om å redegjøre kort for sentrale emner. Slik blir gjennomgangen en generalprøve på muntlig eksamen.

Det som ligger mellom disse ytterpunktene, skal jeg ikke beskrive inngående; det sier seg selv at på så kort tid blir det bare de helt opplagte emner man kan ta opp. Noe som da blir et dilemma, er ønsket om på den ene siden å innføre studentene i leksikografisk teori og på den annen side å gjøre dem kjent med flest mulig typer av ordbøker, bl.a. ved å bla i dem sammen. Det blir alltid for liten tid til det siste.

Et annet dilemma er at vi skal gi en innføring i bruken av elektronisk databehandling (edb) i leksikografien. De fleste har ingen erfaring i bruk av datamaskin; de som eventuelt har det, har sin erfaring fra andre områder enn vårt. Det blir derfor bare en ganske elementær

innføring fra bunnen av, og den må komme såpass langt ute i kurset at man har et grunnlag for å bruke den til noe. Datateknikken er jo ikke noe mål, men et hjelpemiddel. Nå bruker vi to av de tolv ukene til dette.

Etter at Bo Svenséns *Handbok i leksikografi* kom i 1987, finnes det for første gang en utførlig innføring i faget på et nordisk språk. Det er en stor hjelp, og mye av boka er obligatorisk hos oss. Svenséns bok er kjærkommen også fordi pensum ellers er så fragmentarisk, oppsplittet i en mengde artikler eller enkeltkapitler av større verker; flertallet av disse er dessuten skrevet på ikke-nordiske språk og med andre enn disse som emne. Det er ikke skrevet mye leksikografisk teori på norsk. Selv et par av de bedre ting som finnes, er skrevet på engelsk, såsom Finngeir Hiorths "Arrangement of meaning in lexicography" og Trygve Knudsen og Alf Sommerfelts "Principles of unilingual dictionary definitions". Selv med Svenséns utmerkede bok er det på litt lengre sikt bruk for et eget kompendium i tillegg, skreddersydd for våre problemstillinger.

På grunn av alt artikkelstoffet, særlig den valgfrie delen av pensum, ser studieplanen formidabel ut; den valgfrie delen kommer opp i nærmere 70 titler, den obligatoriske i ca. 30. Meningen er selvsagt å tilgodese individuelle interesser hos studentene, f.eks. når det gjelder fremmedspråk. Men det må innrømmes at for pensum som helhet har de fleste en tendens til å gå minst mulig utenom det som er blitt forelest. Valgfriheten blir dermed mye av en helgardering; den blir lite utnyttet, men er der for dem som ønsker å bruke den.

At studentene sverger til magisterens ord når de velger pensum, henger også sammen med behovet for å tilpasse fremmedspråklig faglitteratur til norske problemstillinger. Jeg skal gi noen eksempler på dette.

- I artikkelen "Fremdwortbegriff und Fremdwörterbücher" peker W. Müller bl.a. på systemfremmede bokstavkombinasjoner. I tysk oppfattes dette som et av de momentene som taler for å lansere et hjemlig ord som alternativ, slik tysk jo er kjent for. Også i norsk er dette et moment, men ikke så sterkt, idet vi har en tradisjon for å fornorske skrivemåten, men beholde ordet.
- I artikkelen "Collocability Restrictions in Dictionaries" skjelner E. Aisenstadt mellom *idiomer* og *ikke-idiomatiske fraser*; de siste er igjen inndelt i *frie fraser* og *restringerte kollokasjoner*. Denne bruken av "frase" må nærmest regnes som det en hos oss kaller et syntagme, mens "frase" hos oss er en ganske vid betegnelse for en fast fler-ords forbindelse.
- I klassikeren *Leksikologi og leksikografi* utvider Kristen Møller begrepet "tospråklig" til å omfatte en ordbok på morsmålet når oppslagsordene ikke hører til "det almindelige synkroniske dagligsprog", f.eks. ordbøker over foreldede ord, fagspråk eller dialekter. Dette er et praktisk synspunkt, men ganske ubrukkelig på moderne dialektordbøker eller f.eks. på en ordbok mellom bokmål og nynorsk.
- På grunnlag av engelsk hevder Ali M. Al-Kasimi i *Linguistics and Bilingual Dictionaries* at de fleste ordbøker gir lite av grammatiske opplysninger. Dette stemmer ikke for norsk, idet våre allmennspråklige ordbøker i alle fall er rettskrivningsordbøker - uansett hva de ellers er - og *bøyning* blir behandlet som en del av rettskrivningen og derfor tatt med. "Grammatikk" i norske ordbøker betyr først og fremst bøyning.

- I språk med mindre valgfrihet enn norsk betyr "grammatikk" i ordbøker oftere syntaks, slik f.eks. A.P. Cowie bruker det i "On Specifying Grammar". I sin artikkel slår Cowie sammensetninger og fraser sammen i ett punkt. Det lar seg gjøre i engelsk, der løse sammensetninger er en stor og uproblematisk kategori. Det er den ikke i nordisk; den rommer f.eks. en problematisk form for anglisismer og må iallfall behandles for seg.
- Ordet *dictionary* i engelskspråklig litteratur har ingen entydig ekvivalent på norsk, der det nærmeste riktignok er *ordbok*, men dette må til dels ses atskilt fra *ordliste*; videre inneholder våre *konversasjonsleksika* ofte en større ordbokskomponent (f.eks. definerte fremmedord) enn en engelsk *encyclopaedia* gjør.

En erfaring som sikkert alle har gjort, er at studenter ønsker klare linjer og tydelige grenser. De kan ikke fordra det når samme fenomen kan ses fra mange synsvinkler eller analyseres på flere måter og derfor med flere navn. I forbindelse med ordforrådet er det nødvendig å ta fra dem troen på at et adverbial alltid er et adverb, eller at det finnes fem uselvstendige verb. Det har de bare godt av; det åpner for en videre erkjennelse av virkeligheten.

Verre er det å avfinne seg med at ulike forfattere bruker ulike termer om samme fenomen, eller samme term med ulikt innhold. Det er vi jo vant til, f.eks. i grammatikken, der betegnelser på de ulike aksjonsarter kan bli en terminologisk jungel. Men det er brysomt, og for ikke å gjøre tilegnelsen mer komplisert enn nødvendig har jeg utarbeidet en foreløpig versjon av en leksikografisk termliste på helt elementært nivå, foreløpig på 154 nummer, altså noe større enn Jennifer Robinsons på 135 i tidsskriftet *Dictionaries* 1983 nr. 5.

Det jeg imidlertid har tenkt å ta opp under dette punktet, er en del emner som, så langt jeg er orientert i litteraturen, er uklart avgrenset, byr på uløste problemer og derfor er vanskelige å presentere på en ryddig måte.

Et passende sted å starte kan være med forholdet mellom *ord* og *begrep*. Forskjellen mellom disse står uklart for de fleste. I trykte kilder som vi henter materiale fra, kan vi finne eksempler som "Beatles er et *begrep* i popmusikk" eller "Forbrukerne blir lurt av *begrepet* flat rente". I det første tilfellet må en gjøre det klart at noe konkret, dvs. synlig, ikke er et begrep; hvis det ser slik ut, er det metaspråklig bruk. I det andre eksemplet møter vi forvekslingen av begrep og uttrykk; det er *uttrykket* "flat rente" som lurer forbrukerne. I NORDTERM 1989 finner vi dette:

Begreppen används för att strukturera vetande. Ett begrepp måste inte nödvändigtvis ha något motsvarande språkligt uttryck.

Hvis noe er et aksiom i vårt fag, er det vel at alt kan kalles noe. Det har alltid gjort et særskilt inntrykk på meg at heksene i *Macbeth* utfører "a deed without a name". Det samme ser det ut til at terminologer kan gjøre, mens leksikografer er hjelpeløse "without a name". En må altså skjelle mellom terminologisk og leksikalsk informasjon.

Et problem ved selve ordet er som kjent å avgjøre om det er ett polysem eller flere homonymer. Homonymproblemet har vært angrepet fra mange kanter; det finnes en utilfredsstillende felteori og en do. valensteori, og redaktører har vært mer eller mindre restriktive eller liberale med å utnevne homonymer. På studentene virker det utilfredsstillende at et så viktig område er en stor gråsoner der man redigerer uten faste kriterier, etter hva man

selv finner praktisk. Om jeg skal forsøke å gi kriterier, som jeg gjør nedenfor, må jeg samtidig spørre om flere enn de tre første egentlig er sikre:

- 1 Ulik ordklasse: *sky* subst/adj./verb
- 2 Ulik etymologi: *skaffe* få tak i/spise;  
*gregoriansk*: *g. sang* (av pave Gregor den store, d. 604)/*g. antenne*  
(J. Gregory, skotsk astronom, d. 1675)
- 3 Ulik bøyningskategori: *brenne* sterkt/svakt;  
*la* inf./pret. av *legge*; *nettene* av *natt*/av *nett*; *løvetann*, pl. *-tenner* (på løve)/  
*-tanner* (plante)
- 4 Ulik valgfrihet: *gard* (i *skog*)/*gard* el. *gård*;  
*skap*, m. el. n (fasong)/*skap*, n (møbel)
- 5 Ulikt genus: *skill* m. (i hår)/n. (i vev)
- 6 Funksjonelle restriksjoner: *ligne/likne*,  
passivvendt: *mann og kone lignedes hver for seg*/  
*?faren lignedes av sønnen*
- 7 Ulikt tonem: *for'ball* (av eng. *football*)/*for'ball*  
(muskulatur under fotbladet)
- 8 Total ulikhet i forestilling: *dødsfall* (*døds-fall*)  
fall som fører til døden/*dødsfall* død, bortgang
- 9 Praktiske hensyn, f.eks. stor ulikhet i betydning:  
*?blad*: på tre; publikasjon; skjærende del av redskap...

Særlig 4 er nok en norsk spesialitet, der et ord som helt ut er et polysem i nynorsk, blir et homonympar i bokmål.

Hvis en ser videre på de mindre sikre, fra 4 og utover, oppstår det nye problemer, idet det er lett å se at ordparet i utgangspunktet var ett polysem. En har egentlig ikke løst problemet: Hvor stor avstand mellom betydninger skal det til før polysemet er blitt homonymer?

Vi går over til problemer med den enkelte ordartikkel. Vi tenker oss et oppslagsord med et ganske omfattende materiale, slik at det må bli flere betydningsnumre og en videre oppdeling under dem. Hvordan stiller faglitteraturen seg til dette? Ja, hoveddefinisjonene peker ut *denotasjoner*. De skal samle flere bruksmåter, kjente så vel som mulige, og må derfor være rommelige, også kalt "sterke", nok til dette, sterkere enn det som svarer til betydningen av ordet i den enkelte bruksmåte. I jussen, der en også er opptatt av ordenes betydning, snakker en gjerne om en *kjernesone* og en *randson*. For oss er det vel vanligere å tenke på denotasjonen som kjerne, og rundt denne et forråd av enkelt-bruksmåter eller *konnotasjoner*. Vi ser hvordan enkelte konnotasjoner prøver å tøyne grensen for denotasjonen, f.eks. når et dataanlegg eller en maskin omtales som "skreddersydd". Konnotasjonene peker utover fra sentrum.

Nå peker Ayto (1983) på at konnotasjonene er med og avgjør hvordan den enkelte betydning skal defineres, de bidrar til denotasjonen; sett fra denne synsvinkelen finnes det ingen konnotasjoner, sier Ayto, og tilføyer riktignok at det er en påstand som er åpen for angrep. Jeg vil heller spørre om ikke ethvert betydningsnummer større enn 1 samler et sett av forhenværende konnotasjoner. De peker i samme retning og har sprengt grensen for sin opprinnelige denotasjon.

Selv om vi samler et ganske stort materiale, synes det likevel å bero på redaktørens skjønn når dette har skjedd, og f.eks. *skreddersy* skal få et nytt betydningsnummer. For å ta et eksempel fra *Bokmålsordboka*:

*blinkskudd* skudd som treffer blinken (I,2); overf.: *gaven*  
 var et b-  
*fulltreffer* 1 prosjektil, slag som treffer blink 2 klaff,  
 suksess: *boka, gaven ble en f-*

Redigeringen er ikke feil, men heller ikke udiskutabel: Skulle *blinkskudd* vært inndelt i 1 og 2 eller *fulltreffer* bare hatt én hovedbetydning?

Vi kommer til utformingen av definisjonen og møter det velkjente *substitusjonsprinsippet*, nevnt bl.a. av Knudsen og Sommerfelt (1958); de sier at det stammer fra Hjelmlevs *Prolegomena*. Det går som kjent ut på at definisjonen bør kunne erstatte oppslagsordet i en vanlig tekst. Deres eget verk, *Norsk riksmålsordbok*, er det beste eksempel på at dette ikke kan godtas uten forbehold. Deres enormt lange og delvis anatomiske definisjon av verbet *le* kan på ingen måte ta *le-s* plass uten fullt sammenbrudd i setningen:

*le...1) a)* frembringe korte, ensartede, hurtige på hinannen følgende lyd under støtvide utåndinger (og med lignende ansiktsuttrykk som ved smil), fremkalt ved opplevelse av noget morsomt, komisk, og oftest som uttrykk for glede, munterhet (ofte også for overlegenhet, hån, forakt, ell. for å dekke forlegenhet) ell. ved fysisk påvirkning (purring av hudnervene på visse steder av legemet):...

Det problemet dette skaper, kan det vel likevel finnes en løsning på: Man må gå ut fra den delen av definisjonen som dekker det enkelttilfellet man står overfor i teksten. Det er denne delen som helst skal kunne tre inn i oppslagsordets sted.

Så kommer vi til rekkefølgen av betydningsnumrene. Vi kan f.eks. skjelne mellom en *historisk*, en *logisk* og en *empirisk* rekkefølge. Den historiske, fra eldst til yngst, bryter sammen fordi den ikke lar seg bruke på alle ord. Den logiske skal føre fra én betydning til den neste. Under *barn* har f.eks. *Bokmålsordboka* først "avkom i første ledd etter mennesker", og først som nr. 5 "i sms: *helleb-, lakseb-*". Det er mulig at en mer logisk rekkefølge ville vært å flytte 5 til 2, så man først hadde avkom etter mennesker og derpå etter dyr.

Den tredje inndelingsmåten, den empiriske, er gjerne den man ender med, med et visst hensyn til en av eller begge de andre der det går for det samme og ikke kan misforstås. I sin artikkel fra 1955 redegjør Finngjør Hiorth for denne, med Casares' spanske innføring i leksikografi fra 1950 som eksempel: Først betydninger som svarer til gjengs og vanlig bruk, så de som er foreldede, regionale eller "colloquial", og til slutt tekniske.

*Fraseologien* er et viktig og vanskelig emne i ordboksredigering. Jeg skal ikke gå inn på slike spørsmål som hvordan man velger oppslagsord for å sikre seg at brukeren slår opp riktig. Heller ikke tar jeg stilling til om en frase kan bestå av ett ord, som noen hevder, eller om det må være flere, slik M.J. Wallace (1979) og vel de fleste hevder. Videre finnes det en rekke ulike betegnelser på emnet - frase, fast uttrykk, talemåte, kollokasjon, dels som navn på det samme, dels på grunnlag av ulik inndeling eller ulike synsmåter. Klappenbach (1980) prøver å bringe orden i kaoset ved bl.a. å innføre termen "frasem" om ordene som inngår i

frasen; frasen kan utvides, f.eks. fra "et egyptisk mørke" til "i et *slikt* egyptisk mørke", men frasemene skal ikke være utskiftbare. Det er de imidlertid ofte i praksis, særlig i sine enkelte ledd, fra "skatteskrue" til "*prisskrue*" eller fra "dansen om gullkalven" til "dansen om *oljekalven*". Man kan selvsagt diskutere om slike endringer er "riktige" og f.eks. si at en metafor bør ha en motsvarighet i virkeligheten, og der finnes det ingen "oljekalv"; derfor kan ikke uttrykket godtas. Men ofte må en si at utvidelsen, omskrivningen, allusjonen osv. oppfyller alle krav.

Allén (1976) gir en nyttig inndeling av kollokasjoner i *kombinasjoner*, *konstruksjoner* og *idiomer*. Kombinasjonene viser seg ved frekvens, og de andre to er delmengder av disse: begge typer er grammatisk velformede, men konstruksjonen betyr det som summen av ordene tilsier, mens idiomet ikke gjør det. "Trampe i gulvet" er vel en konstruksjon, "trampe i klaveret" er et sikkert idiom.

Men hva med *metaforen*? Den er et språklig bilde der betydningen på overflaten er den som ordene tilsier, men dens funksjon er å være en tydeliggjørende sammenligning, en *modell*, idet tydeliggjøringen kommer av at den framhever et hovedpoeng ved den virkeligheten den sammenlignes med. Et vanskelig spørsmål er forholdet, og skillet, mellom idiom og metafor. Jeg vil hevde at ethvert idiom har en fortid som metafor, men likevel er problemet der hvis ordboksredaktøren f.eks. vil bruke de to markørene "overført" og "billedlig". Er "gå over bekken etter vann" overført eller billedlig? Jeg kaller det billedlig og en metafor. Skillet bør være at dersom ordet gir riktig mening ut fra sin definisjon, om enn i videste betydning, da er det overført bruk; hvis ikke er det et bilde. I setningen "Dette er realiteter og ikke *kosmetikk*" er *kosmetikk* brukt overført; det er riktignok ikke pudder eller leppestift, men det er iallfall "forskjønnende middel uten reell virkning". Hvis man derimot omtaler sitt lokalmiljø som en "andedam", vil jeg kalle det billedlig, for uansett definisjon kan betydningen "trangsynt miljø" ikke plasseres under samme betydningsnummer som "dam for ender", og en definisjon "lite miljø" blir for "sterk" til at en underbetydning "dam for ender" kan utledes av den. Men i mange tilfeller er skillet langt mer uklart og må overlates til redaktørens skjønn.

Dette var noen eksempler på leksikografiens ulne kanter, dens "fuzzy edges", og siden leksikografistudentene gjerne er motiverte og våkne, har de ingen problemer med å peke ut disse og flere til.

### Hva synes studentene?

Jeg skal nevne noen reaksjoner fra studentene, særlig fra den ene som skrev sin semesterrapport i form av en evaluering.

Han starter med at navnet leksikografi kanskje er uegnet til å vise hvor aktuelt emnet er for studenter i alle språkfag, og at de med et mer opplysende navn kanskje lettere ville ha nærmet seg det. Han ville ha foretrukket en avsluttende *eksamen*, skriftlig og muntlig og med karakter som viste hvilket utbytte man hadde hatt av kurset.

Det kunne vært satt av mer tid i starten til prøver på hvordan begynneren bør gå fram for å trenge inn i leksikografiske problemer - "noen håndfaste arbeidsmetodiske punkter", kaller han det.

Han peker på hvor vanskelig det er å sette opp en individuell pensumliste på grunnlag av så mye valgfri litteratur; studieplanen er "skrekkinngydende i sin rikdom". Det burde også vært mer tid til å presentere alternativer og hva hvert av disse innebærer. Han legger atskillig



vekt på at studenten tidligst mulig i semesteret bør få gjort sitt valg og få det klart for seg hvilke konsekvenser valget har, slik at han fra da av kan "se" den linjen han har valgt og kan studere ut fra den.

Det er mulig, sier han, at studentene alt i starten bør få utpekt hvilke emner som ikke er dekket eller er spinkelt dekket, så de selv kan jakte på litteratur om disse.

Slik det er nå, er han redd at studentene ikke makter å gjøre god nok bruk av litteraturen og av utdelt materiale.

Når studieplanen i innledningen peker på betydningen av ordbokskompetanse, synes han det kan bli for snevert; det burde gjøres klarere at leksikografi er en utmerket påbygningsenhet i ethvert fremmedspråksstudium, og kanskje særlig at leksikografi er høyaktuelt i skolen, eller iallfall burde være det. En slik "markedsføring" kunne trekke flere studenter til faget.

Han stiller opp noen momenter som semesteremnet burde legge vekt på å innføre studentene i når det gjelder ordbøker; blant disse nevner jeg to:

- 1 Si ikke bare hvilke ordbøker som finnes, men mer om forskjellen på dem og hvordan man finner fram til den mest tjenlige ut fra forskjellige formål
- 2 Hvordan kan ordbokens "brukerveiledning" absorberes så godt at eleven/studenten "unngår å slepe på tåpelige misforståelser i årevis, unngår å være plaget av barrierer i arbeidet..."?

Om undervisningen nevner han bl.a.:

- 1 Noen av foreleserne har kanskje lagt for mye an på å dekke hele emnet, og kunne i stedet ha valgt noen hovedlinjer eller -spørsmål, slik at det ble mer tid til drøfting og meningsutveksling
- 2 Man kunne arrangere (spise?)pauser der kurslederen kastet fram "et konkret delspørsmål for fri meningsutveksling"
- 3 I begynnelsen av hver ukesamling bør det settes av tid til å oppklare uklare ting fra forrige uke
- 4 "Terminologisk uklarhet og mangfold skaffet meg hele høsten atskillig bry." Han er fornøyd med den term-listen han fikk utdelt, men den er ikke omfattende nok
- 5 I tilfelle ulike skoleretninger innenfor lingvistikken slår ut i måten å redigere ordbøker på, burde dette være påvist i undervisningen

Jeg kommenterer ikke evalueringen, men lar den tale for seg selv som et eksempel på en nybegynners inntrykk. Heller ikke går jeg i detalj om det som studenten var fornøyd med, som faktisk var det meste. Også de andre som har tatt kurset, later til å ha vært mer fornøyd enn misfornøyd. Men knapp tid er et problem.

### Utviklingen videre

Det er ting som tyder på at leksikografi er på vei til å bli lingvistikkens nye store moteemne. Det kan ha noe å gjøre med det økende antallet språklige minoriteter og behovet for fremmedspråklige ordbøker; andre former for internasjonalt samkvem bidrar også; mange føler behov for å effektivisere sin bruk av språket i konkurransen om publikums oppmerksomhet. Ett moment som også bidrar, er at når en bølge av forskere har vært virksomme på ett område, så vil neste bølge velge et annet. I nyere tid har bl.a. fonologi, syntaks og sosiolingvistikk vært ivrig utforsket. Jeg tror det nå er leksikografiens tur. Det flommer innover oss nye artikler, bøker og tidsskrifter, noe godt og noe ikke fullt så godt.

Alt dette må selvsagt prege leksikografien som undervisningsfag. Vi må holde faget å jour, og den raske utviklingen gjør dette til en krevende oppgave. Bare det nye kjempeverket *Wörterbücher* (Hausmann og fl. 1989-91) er nok til å kreve en omfattende revisjon.

Et semesteremne er det laveste trinn i faget. Vi håper etter hvert å kunne bygge det ut til høyere trinn. Vi må også kunne se for oss en videreutdanning f.eks. etter lignende linjer som ERASMUS' planlagte "European Diploma in Lexicography". En slik utvikling må skje etter behov. Behovet er der, og vår oppgave blir å gjøre det synlig og erkjent. Et ledd i det er denne konferansen, den nordiske organisasjonen av leksikografer, og all annen oppmerksomhet som tjener faget. Vi har for lengst forlatt elfenbenstårnet - om vi noen gang har sittet der; det ligger i fagets natur at leksikografer er utadrettede.

### Noter

1. Etter generaliseringen kommer unntakene: til 5: Når oppslagsordet og hovedordet i definisjonen står i et avledningsforhold og hovedordet selv er definert, regnes det som akseptabelt å vise dit: *kondensasjon* det å kondensere (s.d.); til 6: Særlig i mindre (enspråklige) ordbøker blir en del vanlige ord antatt å høre til brukerens forhåndsviten og blir ikke definert; til 7: mange vil regne ulik bestemthet som akseptabelt i artikler av typen [ /-logi, læren om...).

### Litteratur

- Aisenstadt, E. 1979. Collocability restrictions in dictionaries. I: *Dictionaries and their users*. Exeter Linguistic Studies 4, University of Exeter: 71-74. Exeter.
- Al-Kasimi, Ali M. 1977. *Linguistics and Bilingual Dictionaries*. E.J. Brill, Leiden.
- Allén, Sture 1976. On phraseology in lexicography. I: *Cahiers de lexicologie* XXIX nr. 2: 83-90. Besançon.
- Ayto, J.R. 1983. On specifying meaning. I: *Hartmann 1983*: 89-98.
- Cowie, A.P. 1983. On specifying grammar. I: *Hartmann 1983*: 99-107.
- Dictionary Research Centre, Language Centre University of Exeter. Rundskriv av 31.5.1990: ERASMUS Action IV.2 "European Lexicography Diploma". Exeter.

- Gundersen, Dag 1986. Melding av Eva L. Haugen: *A Bibliography of Scandinavian Dictionaries. With an Introduction by Einar Haugen*. Kraus International Publications, White Plains, N.Y. 1986. 387 s. I: *Maal og minne* 1-2: 90-94. Oslo.
- Hartmann, R.R.K. (red.) 1983. *Lexicography: Principles and Practice*. 228 s. Academic Press. London.
- Hausmann, F.J., Oskar Reichmann, H.E. Wiegand og L. Zgusta (red.) 1989, 1990, 1991. *Wörterbücher I-II-III*. 3355 s. Walter de Gruyter, Berlin - New York.
- Hiorth, Finngeir 1955. Arrangement of meaning in lexicography. I: *Lingua* 4:413-424.
- Klappenbach, Ruth 1980. *Studien zur modernen Deutschen Lexikographie*. John Benjamins B.V., Amsterdam, artiklene: Probleme der Phraseologie, s. 196-221 (først utg. 1968); Homonyme oder polyseme Wort?, s. 235-248 (først utg. 1971).
- Knudsen, Trygve, Harald Noreng og Alf Sommerfelt (red.) 1937-1957. *Norsk riksmålsordbok*. 3160 + 4348 sp. Utgitt av Riksmålsvernet. H. Aschehoug & Co. (W.Nygaard). Oslo.
- Knudsen, Trygve og Alf Sommerfelt 1958. Principles of unilingual dictionary definitions. I: *Proceedings of the Eighth International Congress of Linguists*: 92-101. Universitetsforlaget. Oslo.
- Landrø, M.I. og Boye Wangensteen (red.) 1986. *Bokmålsordboka. Definisjons- og rettskrivningsordbok*. 697 s. Universitetsforlaget. Oslo.
- Müller, W. 1976. Fremdwortbegriff und Fremdwörterbuch. I: *Probleme der Lexicologie und Lexicographie*: 211-223. Pädagogischer Verlag Schwann. Düsseldorf.
- Mæhlum, Helge 1990. Svar på en mild og myk oppfordring om å gi En evaluering av kurset høsten 1990 i Leksikografi. Utrykt semesterrapport for semestererme i leksikografi. Universitetet i Oslo.
- Møller, Kristen 1959. *Leksikologi og leksikografi. Udvalg for Folkemaals Publikationer. Serie A nr. 19*. J.H. Schultz Forlag, København.
- NORDTERM 1989. Publication 2: *Terminologins terminologi*. 36 s. Nordterm, Reykjavík - Oslo - Helsingfors - Stockholm - København.
- Robinson, Jennifer 1983. A Glossary of Contemporary English Lexicographical Terminology. I: *Dictionaries. Journal of The Dictionary Society of North America* No. 5: 76-93. Terre Haute, Indiana.
- Svensén, Bo 1987. *Handbok i leksikografi. Principer och metoder i ordboksarbetet*. Esselte Studium/Tekniska Nomenklaturcentralen. Stockholm.
- Utvalget for forskningsspørsmål, Universitetet i Oslo 1989. Vurdering av ordboks- og arkivarbeid i forhold til vitenskapelige kvalifikasjoner. (Utrykt innstilling.) Universitetet i Oslo.
- Wallace, M.J. 1979. What is an idiom? I: *Dictionaries and their users...*: 63-70, se Aisenstadt, E.
- Zgusta, Ladislav 1971. *Manual of Lexicography*. 360 s. Mouton, The Hague/Paris.

## Fra en passiv til en aktiv ordbok Det kombinatoriske aspektet i fokus

I tradisjonelle ordbøker er betydningen som regel dominerende i forhold til andre faktorer, bl.a. av hensyn til kravet om å begrense ordbokens omfang. Ordbokens opplysninger og definisjoner er gjerne knappe og absolutt formulert. Leseren står vanligvis overfor "endelige" forklaringer på enkelte isolerte ord, men har vanskelig for å oppdage helheter, se innbyrdes sammenhenger og trekke egne konklusjoner. I to islandske ordboksverk som nå er under arbeid, blir det lagt vekt på å imøtekomme leserens behov for informasjon om ordenes kombinatoriske natur, både når det gjelder syntaks og ordlagning. Dermed får også semantikken en skarpere konkretisering til nytte for leseren.

### Nytenkning og tradisjon

Begrensningens kunst er en av de vanskeligste oppgavene for dem som lager ordbøker. De må velge og vrake fra den enorme mengde av opplysninger og stoff som kan være aktuell, bestemme forholdet mellom de enkelte avsnitt og avgrense stoffet slik at det blir et innbyrdes samsvar i framstillingen. I senere tid har datateknikken gjort problemet enda tydeligere. Gjennom innføringen av edb har nemlig adgangen til kjempestore ordsamlinger blitt mye lettere. Samtidig har datateknikken også gjort det lettere å få til den avgrensning og det innbyrdes samsvar som er ønskelig. Ved hjelp av edb har det vært mulig å forbedre framstillingen og utførelsen av de store historiske ordboksverkene og finne mange slags uoverensstemmelser og unøyaktigheter som selv ikke de mest oppmerksomme og skadefro lesere hadde oppdaget tidligere (se Weiner 1989). Og vi har sett nye betoningar når det gjelder viktigheten av de enkelte komponenter i ordboka, som bl.a. henger sammen med at man er blitt mer bevisst om muligheten og ønskeligheten av på en organisert måte å benytte den store mengde stoff som kan brukes, og som det kan tas hensyn til ved ordboksarbeidet. I den forbindelse vil jeg vise til den engelske Cobuild-ordboka, der man på forskjellige måter har gått nye veier når det gjelder presentasjon av stoff (se Sinclair 1987).

Men denne utviklingen har også ført til at man i større grad enn før har prøvd å ta hensyn til brukernes forutsetninger og behov, og det er klart at dataføringen av ordbøker og markeringen av innholdet med forskjellige koder har gitt oss uanede muligheter når det gjelder å få tak i de opplysninger som vi er interesserte i, enten det dreier seg om en hendig ordsamling i forbindelse med tekstbehandling, eller en database av største type som f.eks. den dataførte utgaven av Oxfordordboka. På dette området kan man også finne nyheter i de trykte ordbøkene, men det ligger jo i sakens natur at det er vanskeligere å realisere slike ting der, og ikke mulig å ta hensyn til brukernes forutsetninger og behov på samme måte.

Men det er ikke bare bokformen som gjør det vanskelig for oss å gå nye veier når det gjelder stoffvalg og presentasjon. Vi har gamle og gode tradisjoner i ordboksarbeidet, og de

har sikret en stabilitet som de fleste naturlig nok nøler med å forstyrre. Ja, det er i virkeligheten vanskelig å bryte med denne tradisjonen, fordi vi vet at brukerne er svært vanefaste i sitt syn på bruken av ordbøker. Undersøkelser om folks vaner og forventninger når det gjelder bruken av ordbøker bekrefter da også i stor grad den tradisjonelle presentasjonen av innhold og form, og er ingen oppfordring til fornyelse. De som vil gjøre store endringer i det opplegget brukerne er vant til, må derfor være innforstått med nødvendigheten av å få opp interessen og forståelsen for den type ordbok de presenterer. Og på dette området må vi vel kunne forvente gode resultat i framtida, ikke bare på grunn av innsatsen til den enkelte ordbokforfatter, men også på grunn av det som skjer innenfor selve leksikografien, og fordi kontakten mellom forfattere og brukere stadig blir bedre. Men det er uten tvil lettere sagt enn gjort å endre brukernes tradisjonelle syn på ordboka.

Behovet for nye betoningar i framstilling og disposisjon dreier seg først og fremst om å gi brukerne muligheter til å kunne utnytte ordboka på en bedre måte slik at de kan få tak i flere og bedre opplysninger enn de tradisjonelle ordbøker har kunnet tilby. Dette gjelder ikke bare forklaringer av enkelte ord og uttrykk, der har de gamle ordbøkene vanligvis hatt sin styrke. Det kan også være snakk om å få bekreftet ordbruk, få råd og opplysninger om ordbruk eller språkbruken generelt, eller muligheten til å skaffe seg innsyn i hele ordforrådet eller deler av det. Dette har brukerne bare i begrenset omfang hatt muligheter til i tradisjonelle ordbøker.

### Ordforklaringenes dominans

Kravet om å spare på plassen og begrense størrelsen på ordbøkene kan i seg selv virke som en brems på fornyelse i framstilling og disposisjon. Strengt krav på dette området fører som oftest til at ordbeskrivelsene blir knappe og den indre komposisjon uklar, noe som spesielt går ut over større og vanskeligere ord. Dessuten har dette kravet styrket ordforklaringens dominans i vanlige ordbøker og skjøvet andre ting som f.eks. syntaktiske trekk, brukseksempler og ikke minst ordlagingsdelen til side. Når kravet om sparing er stort, blir opplysninger om forhold som f.eks. bøyning, syntaks og stil gjemt bort i mer eller mindre uklare tall- og kodesystem som brukerne ikke har vært særlig villige til å bruke på den måten man håpet. Den semantiske beskrivelsen med sine detaljerte definisjoner og sin nøyaktige analyse har krevd så mye plass at det har vært nødvendig å begrense omfanget av andre forhold til et absolutt minimum. Som oftest har dette vært til stor ulempe for brukerne.

Den betydning og posisjon som ordforklaringen har hatt i vanlige ord, skyldes ikke bare tradisjonen. Den har også i stor grad vært bestemt av hva slags brukere ordboka er beregnet for, og under hvilke omstendigheter den er tenkt brukt. Man har gått ut fra at det første som skjer, er at brukeren fatter interesse for et spesielt ord. Når ordbokas rolle så er å presentere uttømmende opplysninger om ordet som et avgrenset hele, er det selvfølgelig klart at semantiske forklaringer og semantisk analyse er på sin plass. Behovet for ordforklaringer er selvfølgelig størst hos den brukeren som ikke har oppslagsordets språk som morsmål; etter hvert som brukeren får bedre kunnskaper om språket og lærer å beherske det bedre, vil dette behovet vanligvis minke. Også når det gjelder ordforrådet, kan behovet for nøyaktige ordforklaringer variere. Selv om det er vanskelig å generalisere kan man vel si at behovet er størst når det dreier seg om faguttrykk som f.eks. uttrykk og fenomen i teknikk og vitenskap. Men behovet for ordforklaringer kan også være stort ved uttrykk som ikke hører til vanlig språkbruk. Jo nærmere uttrykkene ligger vanlig språkbruk, jo mindre behov er det for ordforklaringer. Og svært ofte er det slik at behovet er omvendt proposjonalt med redaktørens

problemer og besvær. Den semantiske beskrivelsen av grunnordene, f.eks. beskrivelsen av de vanligste verb og preposisjoner, er som kjent en av de vanskeligste oppgavene for en ordboksredaktør.

### Aktiv bruk

Det vi har i tankene, er å kunne nytte ordboka på en bedre måte og legge mer vekt på den aktive og bevisste brukers forutsetninger og behov. Hvis vi forandrer disposisjonen og reviderer forholdet mellom de forskjellige deler av ord vil det kunne åpne for betydelige muligheter til fornyelse til fordel for brukeren. Den mest interessante muligheten består i å redusere omfanget av den semantiske delen og legge mer vekt på andre forhold. I en slik endring av betoningen ligger det selvsagt ingen dom over ordforklaringens betydning, men det er klart at en lang og inngrodd tradisjon for at det semantiske skal være tyngdepunktet i en ordboksbeskrivelse, har gitt oss økte muligheter og økt frihet til å gå nye veier både når det gjelder disposisjonen og avgrensingen av stoffet. Vi kan nemlig gå ut fra at de brukere som har behov for vanlige ordforklaringer, lett kan finne de rette opplysninger i de tradisjonelle ordboksbeskrivelsene. Og det er virkelig behov for å lage nye ordbøker for å fylle opp det som mangler eller ikke kommer tydelig nok til uttrykk i de ordbøker som finnes. Men selvsagt skal en ny ordboksutgivelse ikke ha som målsetting å gjøre tidligere ordbøker ubrukelige. En slik holdning kan nemlig drepe all originalitet og føre til at den nye ordboka blir en klosset og til og med taktløs etterligning av de verkene som er brukt som kilder.

Endringer i forholdet mellom de ulike deler av ordforklaringen skal ha til hensikt å gi brukerne lettere adgang til opplysningene i ordboka. De skal videre fokusere på de delene som ofte er blitt trent til side og som jeg har nevnt tidligere. For å belyse dette nærmere vil jeg nevne det jeg mener man bør ha som rettesnor i et slikt arbeid:

- \* Det bør legges vekt på å vise at oppslagsordene er en beslektet helhet som gjenspeiler ordlagingens mangfold og produktivitet
- \* Ved ordbeskrivelsen bør det legges mer vekt på å vise brukseksempel enn på definisjoner og forklaringer
- \* Det bør legges mer vekt på å vise en levende og representativ språkbruk enn på det etymologiske og språkhistoriske aspektet
- \* Det bør legges vekt på ordboka som et hjelpemiddel ved praktisk og skapende språkbruk
- \* Språkbrukerens kunnskaper og erfaringsbakgrunn bør få komme til sin rett ved bruken av ordboka

Det som framfor alt betyr noe for resultatet av disse endringene, er at det lykkes å vekke brukernes interesse, pense dem inn på stoffet, overbevise dem om at de virkelig har noe å hente i ordboka, noe som befester den innsikt de allerede har, og som framfor alt tilfredsstiller deres nysgjerrighet, overrasker og gjør bruken av ordboka til en spennende oppgave.

Jeg går ut fra at betingelsene for slike betoningseendringer vil være forskjellige i de ulike språksamfunn. Jeg er heller ikke i tvil om at de som tar på seg ansvaret for ordboksarbeid og har muligheter til å sette sitt preg på det, ikke er like interessert og har ulike syn på nødvendigheten av slike endringer.

## Islandsk ordbokstradisjon

Det er en rimelig forutsetning at forsøksvirksomhet i forbindelse med ordboksarbeid helst bør skje i de språk som har lengst tradisjon med de fleste og mest varierte ordbøker. Der er det også enklest å måle markedets reaksjoner. Det er naturlig at i diskusjonen om ordbøker og ordboksarbeid er det de store språkene - og da spesielt engelsk - og deres forutsetninger som blir lagt til grunn. Den engelske ordboksverden kjennetegnes da også av et antall og en mangfoldighet som langt overgår det vi kan skryte av i de mindre språk som f.eks. de nordiske.

Men ved nærmere ettertanke er dette et synspunkt som slett ikke er berettiget når det gjelder de mindre språkene. Forholdene der kan i virkeligheten gi større muligheter for fornyelse enn i de større språksamfunnene. Grunnen til det er ganske enkelt at en fattigslig og kanskje klart foreldet tradisjon krever fornyelse. Og vi må heller ikke se bort fra at vurderinger som tar utgangspunkt i engelsk (eller for så vidt i et hvilket som helst språk), ikke uten videre kan overføres til andre språk med ulik struktur og språkkultur.

Islandsk har ikke noen lang ordbokstradisjon. De ordbøker som det er verdt å merke seg, er på mange måter både foreldede og mangelfulle når det gjelder moderne islandsk. Det gjelder Sigfús Blöndals *Íslensk-dönsk orðabók* og Árni Böðvarssons *Íslensk orðabók*. Blöndals bok er fra årene 1920-1924, mens *Íslensk orðabók* ikke kom ut før i 1963. I 1983 kom en ny og utvidet utgave av denne ordboka. Disposisjonen av *Íslensk orðabók* bærer i stor grad preg av å bygge på ordboka til Blöndal og at hovedmodellen er en tospråklig ordbok.

Den tradisjonen som gjenspeiler seg i *Íslensk orðabók*, har en del typiske trekk (se Jón Hilmar Jónsson 1985):

- \* Det etymologiske og språkhistoriske aspekt preger i stor grad ordvalg og disposisjon (f.eks. rekkefølgen og forholdet mellom de forskjellige avsnittene i ordforklaringen).
- \* Det er lagt vekt på å bevare den ordbruken som er gammel, dialektpreget og knyttet til kulturtradisjonen (ord, ordforbindelser, betydninger).
- \* Det blir lagt vekt på grunnord (usammensatte ord). Omfanget av sammensatte ord, spesielt de som består av mange ledd, er begrenset.
- \* Holdningen til ordforrådet er normativ. Ved en mengde fremmedord og lånord og betydningslån finner vi spesielle advarsler.
- \* Det semantiske er hovedkriteriet ved disposisjonen.
- \* Ordforklaringene er korte, og består for det meste av synonymer.
- \* Beskrivelsen av den syntaktiske konteksten er knapp og usystematisk.
- \* Brukseksemplene er få og korte.

Jeg tror forresten at disse trekkene i større eller mindre grad er typiske for ordbokstradisjonen i andre språk også, f.eks. de nordiske. Selv om en ordbok av denne typen kan inneholde en mengde forskjellige opplysninger, som f.eks. *Íslensk orðabók* gjør, så er den framstilt på en uklar måte og er vanskelig tilgjengelig for brukerne. Stoffutvalget gir heller ikke et godt bilde av vanlig språkbruk.

Jeg vil særlig nevne to forhold som det bør tas spesielt hensyn til ved en leksikografisk beskrivelse av islandsk. For det første er det det store antall nylaginger som skjer ved sammensetninger og avledninger av islandske ordstammer. Årsaken til den store mengden er bl.a. den begrensede tilgangen på utenlandske ord og ordstammer i språket. For det andre er det nødvendig å ta hensyn til det innviklede bøyingssystemet og den fleksible syntaksen som vi har i islandsk. Dette gjelder spesielt ved beskrivelsen av verb der kasussystemets

mangfoldighet kan føre til at vi får mange ulike typer av kasusreksjon og upersonlige setninger.

### Nye betoningar i islandsk ordboksarbeid

Dette er en situasjon som krever revisjon og nytenkning. Og det er nok av oppgaver å ta av. Fremdeles mangler vi en egen ordbok for moderne islandsk. En virkelig historisk ordbok finnes heller ikke. Et slikt verk har riktignok vært under forberedelse lenge, eller helt siden 1940-tallet da *Orðabók Háskólans* (Universitetets leksikografiske institutt) ble opprettet. Gjennom ekserpering av trykte skrifter helt fra reformasjonen og gjennom forespørslar blant brukerne om språkbruken i moderne islandsk har det etter hvert blitt samlet en stor mengde stoff.

I de senere år har vi arbeidet med å registrere dette stoffet, få oversikt over det og bearbeide det med tanke på en ordboksutgivelse. Dette arbeidet og den erfaring som det har gitt oss, er det som framfor alt danner bakgrunnen for mine refleksjoner her. Og det materialet som etter hvert er blitt gjort tilgjengelig for oss, vil kunne danne basis for et moderne ordboksverk av den typen jeg har omtalt ovenfor.

I forbindelse med dette vil jeg nå gå inn på to islandske ordboksarbeider som jeg har beskjeftiget meg med i de senere år. Det første har jeg arbeidet på i samarbeid med mine kolleger ved Leksikografisk institutt, det er en *historisk ordbok over islandske verb* som for det meste er utformet i sine metodiske grunntrekk og vil komme som første del av det store historiske ordboksverket. Det andre ordboksarbeidet, som ennå til en viss grad er på planleggingsstadiet, har jeg arbeidet med på egen hånd. Det er *Íslensk orðtengslabók*, som skal bli en kombinatorisk ordbok over moderne islandsk.

### Historisk ordbok over islandske verb

Jeg vil starte med å peke på det som er nytt ved verbordboka med hensyn til disposisjon og framstilling. Mest iøynefallende er den faste disposisjonen som gjør det lett for brukeren å finne de opplysninger som han er interessert i, og som hjelper ham til å få oversikt over enkelte avsnitt eller ordartikkelen i sin helhet (se nærmere beskrivelse hos Jón Hilmar Jónsson 1990).

Disposisjonen er i store trekk denne (se også tekstprøven i appendiks 1):

- \* syntaktiske element danner rammen der
  - aktiv kommer først, deretter mediopassiv og til slutt de selvstendige partisippene (preteritum og presens partisipp)
  - innenfor aktiv hhv passiv kommer bestemt subjekt (rekkefølgen er N, A, D og G) før ubestemt subjekt
  - innenfor de forskjellige subjektene kommer intransitiv bruk fremst, først uten partikkel, siden med partikkel i alfabetisk orden; deretter kommer objektet (i rekkefølgen N, A, D og G), med skille mellom "ikke-levende", "levende" og refleksivt objekt
- \* betydningsforklaringene er underordnet syntaksen
- \* faste ordforbindelser er ordnet i alfabetisk rekkefølge under de betydninger de hører til (angitt med eller uten betydningsforklaringer)



- \* brukseksemplene er ordnet etter beleggets alder. Ved markeringen av alder er århundret delt i tre

Jeg vil her fokusere på det som direkte kommer til uttrykk i selve ordboksteksten, men bakom utarbeidelsen av den ligger det en grundig analyse av alle brukseksemplenes ulike funksjoner og ikke bare av de forhold som bestemmer disposisjonen av teksten eller kommer fram der.

Når det gjelder formgivingen av ordboksteksten, plasseringen og oppdelingen av den på siden, har vi i det siste sett nyheter som viser seg å ha stor praktisk verdi. Cobuild-ordboka innførte teknikken med å føye til en ekstra spalte i marginen med tilleggsplysninger om det som kommer fram i selve ordartikkelen. Denne metoden har vi utviklet videre i verbordboka, og den blir brukt i enda større grad i den kombinatoriske ordboka. Dette er ikke først og fremst gjort for å gi flere opplysninger, men snarere for å gjøre brukeren oppmerksom på de forskjellige komponentene i ordartikkelen og vise hvordan den er disponert.

Som det delvis framgår av tekstprøven fra verbordboka, har stikkordene i marginen forskjellige funksjoner:

- \* de skal binde sammen beslektede fenomen som disposisjonen skiller fra hverandre (betydninger, varianter av faste ordforbindelser)
- \* de skal vise til belegg som har spesiell betydning med hensyn til bl.a. bøyingsformer, ordspråk, språkbruk, kultur og folketro
- \* de skal vise til forleddene av verb med samme stamme som er ordnet under beskrivelsen av grunnverbet

Den siste funksjonen berører et viktig forhold som har med synet på oppslagsordene og avgrensingen av dem å gjøre. I *Orðabók Háskólans* skriftspråkarkiv finnes det til sammen ca. 16.000 verb (se en statistisk oversikt hos Ásta Svavarsdóttir o.a. (1992)). Til sammenligning kan nevnes at i ordboka til Blöndal er antallet verb bare ca. 7.000. Av dette antallet er ca. 35 % usammensatte og prefiksløse verb som under ellers like forhold er berettiget til en plass blant oppslagsformene. Blant sammensatte verb og de med prefiks er det riktignok mange leksikaliserte ord som er fullt gangbare i ordboka, men det er også et betydelig antall verb med forskjellige løse forledd som f.eks. forskjellige slags forsterkende ledd, som ikke er berettiget til en plass som selvstendige oppslagsformer. På den annen side er det klart at eksempel på slike ord er fullt gangbare som materiale til beskrivelsen av grunnverbet, samtidig som opplysninger om bruken av forleddene kan ha betydning for brukeren av ordboka. Det blir derfor tatt stilling til alle sammensatte verbformer ved beskrivelsen av grunnverbet, og sammensatte verb og aktuelle forledd blir angitt i verbets hode. Sammensatte verb blir føyd til i teksten der det er ønskelig. Stikkord i marginen viser til eksempel på slike verb i teksten. For å gi brukerne en oversikt over de viktigste forleddene og deres utbredelse blir de presentert som selvstendige oppslagsformer med en utvalgt liste over verb som er knyttet til dem. En slik framstilling tar sikte på å knytte sammen beslektede ord i ordboka og gjøre forbindelsene mellom dem tydelige for brukerne.

Til slutt vil jeg nevne noe som virkelig kan bidra til å øke nytteverdien av ordboka. Det kommer riktignok ikke fram i selve teksten. Det er uvanlig og selvsagt omstridt å begrense det historiske ordboksverket til bare å omfatte verbene. Bak den bestemmelsen ligger det imidlertid gode grunner som jeg ikke har tid til å gå inn på her. På den annen side er det sørget for å gi adgang til det materialet i boka som spesielt angår ord fra andre ordklasser. Det er gjort ved å presentere en spesiell liste over de ord (spesielt substantiv) som finnes i faste ordforbindelser, og vise til de verbene som er knyttet til dem. I hovedsak er det

meningen å ta med flest mulig av de faste ordforbindelsene under beskrivelsen av verbene, så her er det snakk om en betydelig mengde stoff som i virkeligheten hører til beskrivelsen av andre ordklasser.

Jeg håper at denne oversikten har vist i hvilken grad vi prioriterer de forhold i verbordboka som skal være spesielt egnet til å åpne og lette brukernes adgang til den mangfoldighet av stoff som en lang og innholdsrik ordbokstekst inneholder.

### **Kombinatorisk ordbok over moderne islandsk**

Når man ikke er bundet av det historiske aspektet, og oppgaven er å beskrive moderne islandsk, er det fristende å gå enda lenger når det gjelder å bygge på brukernes kunnskaper og forutsetninger. Det som kan være aktuelt, er å redusere ordforklaringens betydning og omfang og gi mer plass til andre ting.

I *Íslensk orðtengslabók* blir disse synspunktene lagt til grunn. Som navnet på boka antyder, er oppgaven å beskrive ordforrådet i språket på en slik måte at den innbyrdes sammenhengen mellom ordene kommer til uttrykk. Og i tilknytning til det jeg sa i stad om det islandske språkets særpreg, er det ikke bare snakk om ordforbindelser og setninger, ordlagingen skal også tilgodeses. En ordbok som legger vekt på sammenhengen mellom ordene, er ingen nyhet. Et av de nyeste og kanskje mest kjente verk av den typen er *BBI Combinatory Dictionary*, som kom ut i 1986. Representanter fra det nordiske språkområdet er *Svensk handordbok. Konstruktioner och fraseologi*, som Ture Johannisson og K.G. Ljunggren redigerte, og *Dansk Sprogbrug* av Erik Bruun.

Jeg vil nå gjøre rede for noen av de synspunkt som jeg legger til grunn ved arbeidet med den kombinatoriske ordboka, og prøve å påvise den verdi som en ordbok av denne type har. Men verken tekstprøven (se appendiks 2) eller de eksemplene jeg kommer til å vise, bør betraktes som en fast og endelig beskrivelse av de ord eller ordformer de hører til. Det er fristende å gi en noe fyldigere beskrivelse enn man kan regne med å få plass til i ordboka, der reduksjonen ikke minst vil gå ut over brukseksempler og enkelte sammensetninger i ordlagingsdelen.

Først vil jeg nevne noen allmenne prinsipp:

- \* Ordforrådet i boka er begrenset til de ord som ikke er avhengige av en semantisk eller encyklopedisk definisjon
- \* Beskrivelsen av de enkelte ord sikter inn på typiske setningsomgivelser (deriblant faste ordforbindelser) og sammensetninger og avledninger som viser noe ved grunnordet
- \* Beskrivelsen skal knytte sammen beslektede ord
- \* Semantiske definisjoner blir ikke angitt selv om en inndeling på semantisk grunnlag kan forekomme, til og med henvisninger i form av synonymer
- \* Det blir lagt vekt på å gi brukeren klare eksempler på konkret bruk

### **Ordlagingens betydning**

Jeg vil hovedsakelig snakke om ordlagingskomponenten, som etter mitt syn har større verdi enn det som hittil har kommet til uttrykk i ordbokssammenheng, og som kan ha mange funksjoner, ikke minst for beskrivelsen av et språk som islandsk.

I tekstprøven utgjør ordlagingen et eget avsnitt, der de aktuelle ordlagingsledd står som overskrift i en liste over ord som er typiske for mønsteret. Når det gjelder prefikser, suffikser og andre bundne ordledd ville en slik liste utgjøre selve kjernen av beskrivelsen.

En god del av slike sammensetninger finnes i en tradisjonell semantisk ordbok også. Men der blir adgangen til ordenes slektskap som regel hemmet av en stiv alfabetisk orden. Betoningen av det semantiske aspektet begrenser også adgangen til forholdsvis frie sammensetninger. Når det gjelder slike ord, er de ofte altfor omfattende, fordi betydningen i virkeligheten er kjent eller umiddelbart forståelig for brukeren. Ved å overlate mer til brukernes kunnskaper og sløyfe forklaringene slik som her, får vi plass til flere ord. Resultatet er et klarere bilde av det som skjer i tilknytning til ordlagingen. En annen praktisk nytte av denne disposisjonen, og da tenker jeg spesielt på islandsk, er at den gir et klart bilde av forekomsten av de forskjellige ordformer [stamme, genitiv (entall og flertall)] i sammensetninger.

Som vi ser i tekstprøven (under *bók*), kan det være mest naturlig å la ordmengden under de enkelte ordformene stå i alfabetisk rekkefølge. Men dersom klassifiseringskriteriene er så klare at en inndeling i avsnitt vil være bedre for brukerne enn den alfabetiske rekkefølgen, er det selvsagt best å gjøre det på den måten (se under *borð*).

Det sier seg selv at ordlister over ordlaging sjelden er uttømmende. Dersom det dreier seg om produktiv ordlaging, har det heller ingen hensikt å prøve å nå det målet. Det er ikke unaturlig at de mest produktive ordlagingene blir forsømt her som i andre ordbøker. Spørsmålet er hvor langt man bør gå i å presentere eksempel på aktuelle ord. Det er fristende å gå nokså langt for å kunne belyse det aktuelle oppslagsordet, og ikke minst ordets betydning, bl.a. for å erstatte det som går tapt når ordforklaringen mangler. Men forholdsvis utførlige ordlister, spesielt over oppslagsord som er andreleddet i en sammensetning - dem er det forresten vanskelig å finne i tradisjonelle ordbøker - har ikke minst til hensikt å stimulere brukernes språkbruk og hjelpe dem med f.eks. nylaginger. Her er det spesielle hensyn å ta i det islandske språksamfunnet. La oss se på noen forhold som belyser betydningen av å presentere typiske sammensetninger som en del av ordbeskrivelsen:

**Semantisk konkretisering.** De sammensatte ordene kan i betydelig grad uttrykke ordets semantiske særpreg. Vi kan ta som eksempel ordet *skúr*, der den tradisjonelle ordforklaringen kunne lyde: "liten (og dårlig) bygning, spesielt brukt som (provisorisk) bu". De enkelte komponenter i ordforklaringen gjenspeiles i følgende beskrivelse av ordlagingen:

*skúr*-: -bygging, -dyr, -hurð, -ræfill, -ræksni, -skrifli, -veggur, -þak ...  
 -*skúr*: áhalda-, bárujárms-, beit(n)inga-, beitu-, býl-, bræðslu-, fiskgeymslu-, fiskverkunar-, geymslu-, salt-, smíða-, timbur-, vegavinnu-, verkfæra-, veiði-, vöru- ...

En slik framstilling er spesielt nyttig for dem som allerede kjenner til betydningen i hovedtrekk, men er ute etter en skarpere presisering.

**Forskjell på synonymer.** En beskrivelse av ordlagingen kan være til stor hjelp når man vil skille mellom synonymer. I *Íslensk orðabók* blir *skúr* bl.a. definert ved hjelp av synonymet *kofi* "hytte, koie". Forskjellen på disse to ordene kommer klart til uttrykk når man ser hvilke sammensetninger som er typiske for *kofi* :

*kofa*-: -dyr, -garmur, -gluggi, -gólf, -hreysi, -hró, -hurð, -rúst, -ræfill, -skrifli,

-tóft, -veggur ...

**-kofi:** bjálka-, dala-, eldiviðar-, fjárhús-, gangna-, gangnamanna-, hunda-, hús-, hænsna-, kinda-, leir-, leitarmanna-, moldar-, sel-, skemmu-, smala-, strá-, timbur-, torf-, veiði-, veiðimanna- ...

**Nyanserende forledd.** Ordforklaringer har ofte liten praktisk betydning når det gjelder velkjente fenomen. Eksempel på dette er grunnfargene, svart, hvit, rød osv. På den annen side kan en beskrivelse av sammensetningene med disse ordene ha stor betydning, ikke minst som hjelp til å få fram de ulike nyanser av fargene:

**-raudur:** al-, blá-, bleik-, blóð-, brún-, dimm-, djúp-, dreyr-, dreyra-, dumb-, dökk-, eir-, eld-, fagur-, föl-, gló-, glóð-, gul-, gull-, haust-, há-, inn-, í-, jarp-, karfa-, kopar-, kvöld-, leir-, lif-, ljós-, loga-, mold-, mó-, purpura-, rós-, rúbín-, ryð-, skarlats-, skær-, sól-, sót-, sterk-, svar-, út-, vín- ...

**Nyanserende etterledd.** Vi har mange eksempler på at ord er svært produktive som første ledd i sammensetninger der det siste leddet har en nyanserende funksjon og kan uttrykke holdninger som kan variere fra antipati til sympati. I en semantisk ordbok får vi en sterk begrensning av slike sammensetninger og gjerne et uklart forhold til definisjoner av slike ord. I en kombinatorisk ordbok vil det være mulig å presentere slike ord som en sammenhengende helhet, der konkretiseringen av de nyanserende ordene har større verdi enn en semantisk definisjon ville ha. La oss se på ordet *kerling* "kjerring" i slik sammenheng:

**kerlingar-:** -afmán-, -andskoti-, -angi-, -álft-, -álka-, -bikkja-, -bjálfi-, -bjáni-, -bryðja-, -djöfull-, -dyrgja-, -fífl-, -fjandi-, -flagð-, -grey-, -gribba-, -helvíti-, -herfa-, -hex-, -hlussa-, -horrim-, -hrota-, -hró-, -hænsni-, -kind-, -kvöl-, -norm-, -púta-, -rass-, -rassgat-, -skratti-, -skrifli-, -skrudda-, -skrukka-, -skömm-, -tetur-, -trunta-, -tuska-, -tötur-, -ugla-, -vargur-, -væfla ...

**Produktiviteten i bundne ordledd.** En beskrivelse av ordlagingen har stor betydning når det gjelder produktiviteten i forskjellige bundne ordledd. Den får i alminnelighet liten omtale i tradisjonelle ordbøker, der det ikke gis muligheter til å gi en samlet oversikt. Ordleddet *-menni* henviser bl.a. til personer med hensyn til en bestemt (positiv eller negativ) egenskap:

**-menni:** 1. eðal-, fjör-, glæsi-, góð-, göfug-, ítur-, karl-, kjark-, lipur-, ljúf-, mikil-, nett-, ofur-, príð-, snar-, snotur-, snyrti-, snöfur-, spengil-, stór-, val- ... 2. bleyði-, dusil-, dyndil-, fauta-, fúl-, glæfra-, gufu-, heigul-, hrak-, hroða-, hrotta-, ill-, linju-, lítil-, lubba-, löður-, ofsa-, ofstopa-, ó-, ópokka-, pinkil-, rag-, ribbalda-, rindil-, rolu-, rosa-, rosta-, rudda-, rusta-, ræfil-, skít-, skríl-, sludd-, smá-, sora-, sóða-, sterti-, svaða-, svaðil-, svaka-, svoila-, tudda-, var-, vesal-, væskil-, þræl- ...

**Forsterkende forledd.** De forsterkende forledd, som ofte er nær knyttet til bestemte ord, er svært gode eksempel på at en beskrivelse av ordlagingen kan ha stor betydning for språkbruken. Slike forledd finner vi oftest sammen med adjektiv. La oss se på adjektivene *ríkur* "rik", *fátækur* "fattig", *hungraður* "sulten", *saddur* "mett", *feitur* "fet" og *horaður* "mager", der hvert ord for seg har sine karakteristiske forledd. Som eksempel på substantiv som er sterkt knyttet til spesielle forsterkende forledd tar vi ordet *logn* "vindstille":

-**rikur**: flug-, for-, mold-, sterk-, stór-, vell- ...  
 -**fátækur**: blá-, sár-, sára-, ör- ...  
 -**hungraður**: ban-, dauð-, glór-, sár- ...  
 -**saddur**: belg-, pakk- ...  
 -**feitur**: ak-, bráð-, spik-, stokk-, stríð, svín- ...  
 -**hóraður**: bein-, glit-, grind-, grút-, skin- ...  
 -**logn**: blanka-, blá-, blíða-, blæja-, dúna-, hvíta-, koppa-, kyrra-, renni-, rjóma-, slétta-, spegil-, stafa-, stilli- ...

**Adgang til ordforråd på bestemte områder.** En slik framstilling av ordlagingen gir brukeren muligheter til å skaffe seg adgang til en ordmengde som karakteriserer eller dekker et bestemt område. I en ordbok som er ordnet etter alfabetet, er dette en bruk som ikke ligger klart i dagen. Men dersom brukerne er klar over muligheten, kan de med egen oppfinnsomhet skaffe seg mange slags opplysninger.

Islandske vær- og naturforhold framkaller en mengde ord med forleddet *norðan*- "nord-" som henviser til storm og kulde, mens *sunnan*- "sør-" har en annen karakter:

**norðan-: 1.** -andvari, -áhlaup, -átt, -belgingur, -brim, -bylur, -bál, -él, -fjúk, -garri, -garður, -gjóla, -gjóstur, -gola, -gustur, -hörkur, -hret, -hríð, -hvas sviðri, -kafald, -kaldi, -kalsi, -kast, -kul, -kuldi, -kylja, -kæla, -kólga, -nepja, -næðingur, -rok, -rosi, -sjór, -stormur, -straumur, -strekk ingur, -strengur, -stórhrið, -stórviðri, -svarri, -svelja, -sveljandi, -súgur, -veður, -vindur ...

**sunnan-: 1.** -andi, -andvari, -átt, -blika, -blær, -gola, -hláka, -hvas sviðri, -rigning, -rok, -rosi, -stormur, -straumur, -vindur, -þeyr ...

Navn på karakteristiske naturtrekk som f.eks. *melur* "sandslette" viser stor variasjon i sammensetninger, der man får adgang til navn på forskjellige formasjoner:

**mel-: 1.** -alda, -ás, -bakki, -barð, -brekka, -brún, -bunga, -drag, -eyri, -fláki, -haft, -hjalli, -holt, -horn, -hryggur, -hóll, -hæð ...  
 -**melur**: aur-, ár-, blágrýtis-, bruna-, eyði-, grjót-, hrjóstur-, ísaldar-, jökul-, óræktar-, rauða-, rofa-, sand-, stórgrýtis-, uppblásturs-, urðar-, vikur-, öræfa- ...

### Materiale til beskrivelse av ordlagingen

En beskrivelse av ordlagingen som denne her, som strekker seg langt ut over grensene for det ordforrådet som uten tvil er leksikalisert, er i stor grad avhengig av adgangen til store ordsamlinger. Den dataførte ordsamlingen ved *Orðabók Háskólans* som jeg nevnte tidligere (skriftspråksregisteret), er en svært stor og omfattende samling, selv om den selvfølgelig ikke er tilfredsstillende på alle måter. Materialet der har vist seg godt brukelig ved valget og avgrensningen av de ord som er kommet med i sammensetningsrekkene under de forskjellige oppslagsord. Opplysningene om antall eksempel og tidfesting av belegg gir oss verdifull informasjon om ordenes betydning.

Det vil alltid være delte meninger om hvilke ord som bør tas med i sammensetningsrekkene. Det er heller ikke mulig å få med alle ord som egentlig burde være der. Det er spesielt fire forholdsregler man kan treffe for å sikre at ordene kommer med:

- \* En tradisjonell ekserpering av flere skrifter
- \* Sammenligning med ordbøker og ordsamlinger
- \* Sammenligning med dataførte tekstsamlinger (f.eks. leting etter bestemte ordstammer)
- \* Hjelp av spesialister på forskjellige områder

Det er en skjønnsak og i virkeligheten et spørsmål om tid og besvær hvor langt man egentlig bør gå når det gjelder slike tiltak. Men det er i hvert fall god grunn til å framheve den enestående verdi denne ordsamlingen har for en leksikografisk beskrivelse av denne type.

### Den syntaktiske delen

Når det gjelder de syntaktiske forbindelsene, så stiller ordklassene ikke likt. Selv om det kan se ut som om substantivene er den ordklassen som det er vanskeligst å beskrive på en ordentlig og systematisk måte, må vi ikke glemme den særstilling som de vanligvis har sammenlignet med andre ordklasser. På grunn av brukernes kjennskap til dem fungerer de ofte bra som atkomstvei til ordforbindelsene. Adgangen til de begrep som er formulert i ordforbindelsene "fletta bók" og "kaflí í bók" er lettere dersom *bók* er oppslagsordet enn om de andre ordene er det, fordi *bók* er grunnordet og et grunnbegrep som vi må gå ut fra at brukerne kjenner (se Hausmann 1985). I de tilfeller der det finnes en mengde ubeslektede ordforbindelser, kan det være behov for en bestemt klassifisering. En slik klassifisering kan ta utgangspunkt i andre forhold enn det semantiske. Ved f.eks. ordet *bók* vil en semantisk beskrivelse ikke være til stor hjelp. Da gjelder det å tilpasse klassifiseringen så godt som mulig til de betingelser som gjelder, knytte sammen det som rent begrepsmessig er felles og det som viser til samme funksjon eller synsvinkel. I tekstprøven får vi en slik beskrivelse av ordet *bók* der stikkordene i margin skal vise til stoffet i de forskjellige avsnittene.

I andre tilfelle kan det passe bedre å peke på de semantiske trekkene, enten med frie antydninger om den betydning som er aktuell, eller med henvisninger i form av synonymmer (sml. *borð*).

Selv om substantivenes særpreg gir anledning til ulike framstillinger, kan man likevel tenke seg et felles mønster når det gjelder bestemte forhold. Ofte har substantivenes flertallsform større betydning i språkbruken enn entallsformen, eller den har en betydningsmessig særstilling overfor entallsformen som det kan være grunn til å presisere ved å beskrive den som et underordnet oppslagsord (sml. *bót* - *bætur*).

Som kjent krever verbene en regelmessig og systematisert beskrivelse av de syntaktiske omgivelsene. Det grammatiske systemet i islandsk gjør det nødvendig å ordne stoffet i en fast syntaktisk ramme lik den vi finner i verbordboka. Hovedskiller i beskrivelsen kan gjøres synlige på forskjellige måter. På den ene side er det grunn til å gjøre mediopassiv og selvstendige partisipper til underordnede oppslagsord (sml. pl. av subst. i noen tilfelle). På den annen side er det mulig å bruke marginen til å gjøre rede for forskjellige varianter av kasusreksjon og upersonlige setninger. Der kan man også vise til faste ordforbindelser som på grunn av den faste setningsstrukturen blir spredt omkring i beskrivelsen. Agensen markeres atskilt fra verbdelen (innenfor skarpe parenteser). Den syntaktiske valensen kommer til uttrykk ved hjelp av forkortelsene for de relevante kasusformer. Den semantiske valensen blir

formulert med illustrerende ord eller uttrykk med hakeparentes - og gjerne med en indikasjon om at en slik ordrekke ikke er uttømmende (sml. *borða* og *skapa*).

I beskrivelsen av adjektivene står korrelatets avgrensning i sentrum. Den semantiske valensen blir beskrevet på samme måte som ved beskrivelsen av verbene (med hakeparentes).

Som det har kommet fram i de eksemplene vi har sett på, er det i beskrivelsen av alle ordklassene tatt med illustrerende (konstruerte) brukseksempel som skal belyse den mer formelle beskrivelsen. Det blir lagt vekt på å begrense antallet eksemplene og heller ha dem så fullstendige og klare som mulig.

Jeg har bare så vidt berørt den syntaktiske beskrivelsen. En fullstendig samordning der er ikke mulig før man har behandlet hele ordforrådet og det blir mulig å sammenligne ordene etter behov. Som generelle trekk vil jeg legge vekt på ordklassenes ulike posisjoner og den faste beskrivelsen av verb i motsetning til den friere behandlingen og det friere valget av substantiv. Hva dette angår, kommer adjektivene midt imellom verb og substantiv.

### Det kombinatoriske aspektet i fokus

Jeg har prøvd å vise nytten av en leksikografisk beskrivelse der den semantiske delen har måttet vike for en beskrivelse av syntaks og ordlaging i den delen av ordforrådet der en slik beskrivelse egner seg best.

Selv om det vil være en viss konkurranse mellom de semantiske og de kombinatoriske forhold, så er det ikke dermed sagt at det ene behøver å ødelegge for det andre. I verbordboka er de begge med for fullt. Selv om det semantiske er underordnet, er det ikke firt på kravet om at ordforklaringer skal være med i en historisk ordbok. Og ved hjelp av krysshenvisninger får beslektede forklaringer komme til syne i forskjellige strukturer.

I en kombinatorisk ordbok over samtidsspråket av den typen som jeg har snakket om, er det semantiske aspektet skjøvet enda lenger til side. Mangelen på egentlige definisjoner blir imidlertid til en viss grad kompensert ved beskrivelsen av ordlaging og den syntaktiske konteksten, og ved inndelingskriterier som har med betydning å gjøre. I en ordbok av denne type kan man så avgjort gå enda lenger når det gjelder å vise til betydning, f.eks. med å vise til antonymer. Men det viktigste er at verdien av slike semantiske opplysninger blir betydelig større overfor den strukturelle bakgrunn de har her enn i en tradisjonell semantisk ordbok eller synonymordbok.

Jeg mener at ved å legge vekt på det kombinatoriske aspektet kan man oppnå mye som ellers ville gå tapt dersom det semantiske blir lagt til grunn. En del av dette kom jeg inn på i begynnelsen av artikkelen. Til slutt vil jeg understreke det som har vært målsettingen med dette ordboksarbeidet:

- \* De enkelte oppslagsord er ikke noe isolert hele, men står i tydelig forbindelse med andre ord og andre deler av ordforrådet
- \* Ordenes særpreg (inklusive betydningen) gjenspeiler seg i en konkret sammenheng
- \* Brukerne får tilbud om forskjellige atkomstveier til opplysningene i boka
- \* Nyttens bestemmes framfor alt av brukernes egen aktivitet og oppfinnsomhet

I fellesskap styrker disse forhold den språkpedagogiske funksjon som jeg mener at ordbøker i aller høyeste grad bør ha.

**Litteratur**

- Ásta Svavarsdóttir, Jón Hilmar Jónsson, Kristín Bjarnadóttir. 1992. Fra seddelsamling til database. Leksikografisk analyse af islandske verber. I: Fjeld, Ruth Vatvedt: *Nordiske studier i leksikografi. Rapport fra Nordisk konferanse i leksikografi mai 1992.*
- Hausmann, Franz Josef. 1984. Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. I: Henning Bergenholtz & Joachim Mugdan, red. *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28.-30.6. 1984 : 118-129.* Tübingen.
- Jón Hilmar Jónsson. 1985. Íslensk orðabók handa skólum og almenningi. Ritstjóri: Árni Böðvarsson. 2. utg. Reykjavík, 1983. [Resension.] *Íslenskt mál* 7 :188-207.
- Jón Hilmar Jónsson. 1990. A Standardized Dictionary of Icelandic Verbs. *Papers from the Seventh Scandinavian Conference of Computational Linguistics*: 268-296. Reykjavík.
- Sinclair, J.M., red. 1987. *Looking up. An account of the Cobuild Project in lexical computing.* London.
- Weiner, Edmund S.C. 1989. Editing the OED in the Electronic Age. *Fifth Annual Conference of the UW Centre for the New Oxford English Dictionary. Dictionaries in the Electronic Age. Proceedings of the Conference*: 23-31. Oxford.



## Appendiks 1

## Prøver fra Historisk ordbok over islandske verb

þagga

þamba

**þagga** -aði, -að; niðurþagga  
Sjá Orðabók Fritznars.

**PAGGA:**

**þagga niður í e-:** *líta e-ð þagna e. hljóðna 20f* Sjá, eg hefi sefað sál mína og þaggað niður í henni. Sálm. 131,2 (1908) **20m** Við Kristiansand tókst Þjóðverjum ekki að þagga niður í virkjum þeim, er vörðu innsiglinguna. ÓlHansHeims. I, 66 **20s** En það er engin leið að þagga niður í óánægjuróðinni; hún hefur sinn eigin vilja. TurwSekt., 8 **20s** afi minn gamli skundabi beina leið inn í kór ... þaggaði niður í orgelinu og hóf upp raust sína. EBragAf., 25

**þagga niður í e-:** *líta e-n þagna e. hljóðna 19m* Ríkarður þaggar niðri henni. JÁbj.<sup>2</sup> V, 295 **20f** Af munn barna og brjóstmylkinga hefir þú gjört þér vígi, sakir fjandmanna þinna, til þess að þagga niður í óvinum þínum og fjendum. Sálm. 8,3 (1908) **20m** urraði hann eins og úfur og þaggaði á þann hátt niður í hundunum. KipiDýrh., 309 **20m** Auðunn var þar nærstaddur og þaggaði niður í honum áður en hann fékk sagt meira. Skafþbj., 233 **20s** Varla yrði nú þjóðarsorg þótt þaggað væri niður í eins og einu sýslumannstúri vestur á Fjörðum. RagnArUpp., 34 **20s** Vilhjálmur sælaði að fara að opna munninn en ábótinn þaggaði niður í honum með handarhreyfingu og hélt áfram ræðu sinni. EcoRós., 418

**PAGGA E-Þ:** *kveða e-ð niður; bæla e. lægja e-ð 16s* Ef að eg setta ei Saalu mína / og þaggaðe hana. Sálm. 131, 2 (CP) **19f** gátu Rússar eigi með öllu þaggað [uppreistina]. Skirn. 1832, II **19s** \*skal aldrei krossins þungbær þröng / minn þagga fögnað lengi. Sálmb. 1886, 79 **19s** Þú biðst Indverja beri sig upp undan því að menn ... láti þá þar eystra deyta, þá mætti nú þagga það með lagi. Þjarki. 1897, 151 **19s20f** \*og stákk honum svefnþorn, sem sefar hvert mein, / sem svefir að fullu og þaggar öll kvein. SGStAnd. I, 105 **19s20f** \*Þangað til að þennan klíð / þaggar blundur sætur. ÞÉRit. I, 109 **19s20f** \*Og bumburnar þrýtur að þagga þann óð, / sem þrymur þar tryldast í ranni. ÞÉRit. I, 142 **20f** \*þá kný og þagga eg streng, sem fallinn ómar. EbenLj. III, 40 **20f** strit þitt og stríð er á enda og sorgir þínar þaggaðar. Þórbólik., 136 (1915) **20ms** um hana fór ljúf ró, er þaggaði allan kvíða. KristmGBróð., 181

**þagga e-ð niður kveða e-ð niður; bæla e. lægja e-ð 19f** Það er ómöguligt að þagga sérhverja raust niður. Árm. I, 180 **19m** að þagga niður umræðu um [málin]. NF. VI, 101 **19m** Bóndinn, sem þaggað hafði niður þenna kvitt. JÁbj.<sup>2</sup> III, 364 **19m** Að nokkrum árum liðnum hafði Ottó þaggað allar þær óeirðir niður. MeIMið., 35 **19m** eg skal láta Nebúkadnesar Babels konung þagga niður glauminn í Egyptalandi. Esek. 30, 10 (1841) **19m** og var því fjarri, að hún reyndi að þagga niður ástríðu þessa. Þús. I, 47 **19s** hafði Svefs gert hann að ráðsmanni yfir vindunum til að þagga þá niður eða æsa þá upp eptir vild sinni. StollGod., 91 **20f** til þess að þagga niður í honum vitleysuna. SHeiðSt., 6 **20s** Hljóðdeyfirinn þaggaði niður hvellinn svo að hann varð aðeins hviss. DMorðrób., 215 **20s** Sagði, að hann myndi eflaust ráð til þess að þagga niður í henni málaði. Hamk. 280

**PAGGA E-N:** *líta e-n þagna e. hljóðna 16m* hann þaggaði þá / at þeir segdi þat ei neinum. Mark. 8, 30 (OG) **17f** en mælti nokkur þar í móti, þá þaggaði Christjan þá alla. Safn. I, 95 (E) **17m** Þá svarar skráddarinn og bólvir hinum trúlausna mjög og þaggar hann. JÓlind. II, 136 **19fm** Apenna ... þaggaði lýðinn. II., 33 **19m** hún þaggaði barnið og mælti. JÁbj.<sup>2</sup> V, 381 **19s20f** annar

ræðarinn ... þaggaði okkur með bændingu. EbenLm. I, 218 **20f** \*Landaldan þærir barminn svo hljótt / sem barn sé þaggað og smáhatt að rugga. EbenLj. II, 244

**þagga e-n niður láta e-n þagna; kveða niður-tal, skoðun e. andmæli e-s-** **16s** að þú meiger svo þitt skimp fremja / þat elngenn þagge þig niður. Job. 11, 3 (CP) **16s** þa er mjög auðvelt niður at þagga Motstæðu mennena. Post., Form (CP) **18f** Ef nokkur mælti þessu í móti, þaggaði Christian skrifari þann sama niður með stórum og vondum orðum. JHðisk. I, 160 **19m** þegar þagga á niður þá sem telja í tormerki kostnaðarins. Rvikuþr., 28 (1841) **19m** Og ekki er trútt um, að sumir hér í landi séu með fégjöfum um að þagga niður þá, sem þeir hafa helzt beyg af. Margtsend., 160 (1845) **19s20f** \*Sit ég hér enn, fram á eifasta slaginn, / unz svefninn þaggar mig niður. SGStAnd. II, 364

**PAGGAST:** *hljóðna, kyrrast 20f* \*allt á að þaggest, unz allt er gleymt. EbenLj. I, 236 **20fm** Allt öldurót þaggest, þegar ísinn nálgast firðina. GFRit. VII, 290 **20m** \*Er hafgolan þaggest og sólskin frá sjónum / um Sólbakka er. GIKrSóL., 32 **20m** Óll sálarátök hjaðna og þaggest við hinn stíllta og yfirliætislausna söng vatnsins í læknun. NNjarðv-Senuþj., 12

**þaggest niður; vera kveðinn niður 20f** Óll forlagatru er mesta hjátru, sem á að þaggest niður af öllum skynsömum mönnum. Skirn. 1909, 357

**PAGGANDI:** *sem lætur [e-ð] hljóðna 20fm* \*Brosilt og þaggandi lagnættið leið. TómgudmLj., 209 **20s** \*frá rismiklum pálmum, sem hófust í þaggandi blæ. TómgudmRit. II, 49

**þamba** -aði, -að; sf., smá-, þrá-

**ÞAMBRA:** **Þ1.** *drekka ákafi 18m* \*.Þambið þið stúllkur / af þessari skálu. ÓDávVik., 69 (um 1740) **19s** [ungbarnið] kemst fljótt upp á að ná sjálf í pelann og er þá á milli dúranna að smáþamba úr smá-  
honum. Fjalk. 1887, 20 **20s** Mamma þín brosti og þambaði úr glasinu í einum teyg. GuðArSóla., 98 **2** *ganga af ákæfð, sækja fastgönguna 17* \*Eftir þeysa, þamba og geisa. RjFíór. XV, 129 **17ms** \*Þorskarin þorðu ei doska, / þambaði langan strembin. HPSkv. II, 422 **19s** \*Þeir þeysa og þamba / og þurfa margt til vamba. ÓDávþul., 262 **0** \*Þíum hlúm þamba, / þörmin litu þamba, / þjóðv.  
fram um fjalla kamba / þau fara að leita lamba. EÖSFrgr., 224 **19ms** Hann þambaði allt hvað af tók. Fjótstáshérad

**ÞAMBRA E-Þ:** *drekka e-ð af ákæfð, svolgra e-ð 17* \*Það er holt fyrir þyrstan mann / að þamba kopar heitan. ÓDávþul., 305 (17. öld) **18s** [hann] þambaði fullt glas af sárkoldu vatni. CampeSið., 8 **18s19f** \*sátu þá æ í sömu kvöl / sjóðheitt þambandi píslaröl. Þort. I, 630 **19f** til að þamba [vatn] sér til heilsuþótar. Klp. VIII, 181 **19f** \*Eins ok hafsvelgr / einn í norðri / þráþambar sjó / í þurrar fjörur / öllumegin at. KlopMess., 81 **19m** Þegar hún var þyrst, lét hann hana þamba blátt vatn. JÁbj. II, 355 **19ms** gerðu lituð annað en drífa og þamba bjór. BCRöndRit. IV, 455 **19s** Þeir þamba bjór í pottatali. Þjóð. 41, 169 **19s20f** \*Við liggjum í gasinu græna / og góðviðrið þömbum í teyg. SGStAnd. I, 61 **20f** Kaffi var ekki í þá daga þambað sem nú. Elm. '14, 57 **20fm** síþambandi te, sem þeir fengu ókæypis um borð. BSLL., 212 **20m** Kyllir vömbina af kjöti og þambar vatn á eftir. KipiDýrh., 115 **20s** Hinn fyrrnefndi með séríglas fyrir framan sig en meistari þambandi te. JDanHaf., 16 **20s** Þau borðuðu auðvitað alla afmælistertuna mína og þömbuðu allt gosið. ÞórbHeljLilL., 24 **20s** Fyrir mér var áfengi svartur dauði sem þreyttir og sktugir sjóðar þömbuðu af stút og urðu blíndfullir af. MattVIBStröð., 15

AL

smá-

þjóðv.

AT

þrá-

&lt;yfirt&gt;

sí-

## Appendiks 2

### Prøver fra Íslensk orðtengslabók (Íslensk kombinatorisk ordbok)

#### I. Substantiv

<p>"møbel" "spising" "deler" ="skipsbord" ="side", "lag"</p>	<p><b>borð</b> no-hvk 1. setjast/sitja við borð, breiða/setja dúk á borð(ið); leggja á borð (fyrir e-n); taka af borðinu; dúkað borð; standa upp frá borðum; undir borðum skólastjórinn hélt snjalla ræðu undir borðum; kalt borð það er fastur síður á hótelinu að hafa kalt borð í hádeginu 2. fótur/fætur á borði; plata á borði; brún á borði. 3. vera/fara um borð; fara frá borði; innan borðs skipið lagði að bryggju með forsetann og fíruneyti hans innan borðs; falla fyrir borð 4. efra/neðra/ytra/innra borðið (á e-u[efni, klæði, vegg, þilju]); á ytra borðinu hún getur verið hrjúf og kaldranaleg við fyrstu kynni en það er bara á ytra borðinu 5. borð á e-u[fláti].</p> <p><b>borð-:</b> 1. -brún, -búnaður, -bæn, -dama, -diskur, -dúkur, -félagi, -fótur, -hald, -hera, -hnúfur, -krókur, -lampi, -plata, -ræða, -rönd, -salur, -sálmur, -síður, -silfur, -skreyting, -stofa, -tennis, -vín ...2. -hár, -lágur, -stokkur ...</p> <p><b>borðs-:</b> -endi, -horn ...</p> <p><b>borða-:</b> -röð ...</p> <p><b>-borð:</b> 1. afgreiðslu-, bar-, borðstofu-, eldhús-, hlað-, horn-, jóla-, kaffi-, kjör-, knatt-, lang-, matar-, próf-, skák-, skóla-, skrif-, skurðar-, smíða-, sófa-, spila-, stofu-, sölu-, tafl-, te-, teikni-, veislu-, vinnu- ...2. hljóm-, ljósa-, mæla-, nótna- 3. skjól-</p>
<p>"leser" "forfatter" "oversetter" "utgave" "form" "mottakelse" "innhold" "deler"</p>	<p><b>bók</b> no-kvk 1. lesa bók; lesa í bók; líta/glugga/kfja í bók; taka sér bók í hönd; taka fram bók (úr hillu), fletta bók; blaða í bók; opna/loka bók; liggja í bókum 2. semja/skrifa/rita/taka saman bók; uppkast/drög að bók; senda frá sér (nýja) bók; tileinka e-m bók; árita bók (sína) 3. þýða bók; frumsamin/þýdd bók 4. gefa út bók; ritstýra bók; binda (inn)/prenta bók; útgáfa/prentun bókar; bók kemur út; bók fer í prentun; upplag bókar; (mörg) eintök af bók 5. bók er í stóru/litlu broti; bók er í bandi/innbundin/óinnbundin 6. skemmtileg/spennandi/fróðleg/merkileg/athyglisverð/vönduð/leiðinleg/léleg/litilfjörleg/ómerkileg/...bók; bók er vel/illa skrifuð; gagnrýna bók; bók er vel/illa tekið; bók fær góðar/slæmar viðtök; bók fær góða/slæma dóma; bók er á þrotum; bók er uppseld 7. bók (er/fjallar) um e-ð; bók inniheldur e-ð; bók hefur e-ð að geyma; efni/innihald bókar. 8. blaðsíða/síða/opna/kafli í/úr bók; spjöld/kápa/kjölur/band/letur á bók.</p> <p><b>bók-:</b> -band, -fell, -fræði, -fróður, -færa, -færsla, -hald, -hlaða, -lestur, -mál, -menntir, -nám, -sala, -sali, -stafur, -vit ...</p> <p><b>bókar-:</b> -heiti, -kafli, -kápa, -titill ...</p> <p><b>bóka-:</b> -búð, -flóð, -forlag, -herbergi, -hilla, -hnúfur, -kostur, -maður, -markaður, -merki, -ormur, -safn, -skápur, -stofa, -stoð, -stóll, -uppboð, -útgáfa, -vörður, -þjóð ...</p> <p><b>-bók:</b> ár-, banka-, barna-, bréfa-, dag-, ferða-, fræði-, gerða-, gesta-, hand-, kennslu-, kvæða-, les-, lestrar-, lita-, ljóða-, lög-, matreiðslu-, metsölu-, minnis-, mynda-, orða-, reiknings-, sálma-, skóla-, skrif-, sparisjóðs-, stíla-, sýnis-, sögu-, vasa-, vasabrots-, vísna-, ær- ...</p>

- = "lapp"  
= "helbredelse"  
"ordforb."
- bót** no-kvk 1. setja/líma bót á e-ð[flík, slöngu ...] 2. það er (mikil/lítill/engin) bót að e-u[ráðstöfun, lækningu, meðali]; fá (mikla/nokkra/litla/enga) bót (við e-ð[aðgerð, uppskurð]) 3. e-að[meðal, heilsudrykkur ...] er allra meina bót *anna mín taldi jurtae vera allra meina bót; eiga ekki bót fyrir rassinn á sér hann getur dreiddanlega ekki borgað þér enda á hann ekki bót fyrir rassinn á sér; það er (þó) bót í máli (að ...) það var slæmt að tapa veskinu en það er bót í máli að það var tómt; lofa (e-m) bót og betrun; mæla e-u[hægðun, framferði, ástandi]/e-m (ekki) bót engum dettur í hug að mæla gyðingaofsóknun nasista bót; ráða bót á e-u[slæmu ástandi, vanda, erfiðleikum] stjórnnni gengur illa að ráða bót á arvinnuleysinu.*
- bót: á-, bragar-, bú-, endur-, harma-, heilsu-, meina-, nafn-, rauna-, réttar-, sagn-, sálu-, sára-, sið-, skinn-, stjórnar-, til-, upp-, við-, þokka- ...
- "for skade/tap"  
"ordforb."
- bætur** flt 1. eiga rétt á bótum (fyrir e-ð[tjón/skaða]/eftir e-n[maka]); fara fram á bætur (fyrir e-ð); krefjast bóta fyrir e-ð; fá (greiddar) bætur (fyrir e-ð/efitir e-n[maka]); greiða/borga (e-m) bætur 2. bíða þess ekki/aldrer bætur (að ...) *hún slasðist svo illa að ég er hræddur um að hún bíði þess aldrei bætur; e-að[ráðstöfun, aðgerð] er til (mikilla/stór-) bóta fiskifræðingar telja að það yrði til mikilla bóta að takmarka veiðarnar; e-að[það/þetta] stendur til bóta það er ekki enn búið að gera við girðinguna en það stendur til bóta.*
- bóta-**: -ábyrgð, -fé, -fjárhæð, -greiðsla, -hækkun, -kerfi, -laus, -réttur, -skýrsla, -tímabil, -þegi ...
- bætur: atvinnuleysis-, barna-, endur-, fé-, hafnar-, hags-, húsnæðis-, jarða(r)-, kaup-, kjara-, kyn-, laga-, launa-, lífskjara-, mann-, máls-, miska-, ó-, réttar-, sára-, skaða-, stjórnar-, tún-, um-, úr-, vaxta-, veg-, verð-, yfir- ...

## II. Verb

- "m. partikkel"  
("m. akk.")
- borga** so <e-r> 1. borga af e-u[lán, skuld] → afborgun; borga fyrir e-n *leyfðu mér að borga fyrir okkur bæði; borga fyrir sig ég legg til að hver borgi fyrir sig, hann borgaði fyrir sig í hvert skipti sem þeir reyndu að hrekja hann; borga (e-ð[upphæð]) inn á e-ð[kaup; reikning] ég get borgað inn á þetta en ég get ekki borgað þetta út → innborgun; borga (e-ð[upphæð]) með e-m[barni, skjólstæðingi] hann borgar með börnunum sem hann á með fyrri konunni; borga með sér þeir sem fá inni á elliheimilinu verða að borga með sér; borga (e-ð[upphæð]) undir e-ð[bréf, pakka, sendingu]; <e-r/e-að[fyrirtæki, stofnun]> borga út það er yfirleitt borgað út á föstudögum fá borgað út ég fæ ekki borgað út fyrr en eftir helgi → útborgun 2. <e-r/e-að[stofnun, fyrirtæki]> borga e-ð[upphæð; vöru; skuld, lán, reikning], borga (e-ð) í peningum/með ávísun/með greiðslukorti; fá e-ð[hlut, vöru, skuld] borgað heldurðu að þú fáið blínn nokkurn tíma borgaðan?; þú skalt/hann/...skal fá þetta borgað ég læt hann ekki koma svona fram við þig, hann skal einhverntíma fá þetta borgað; borga brúsann ofi verður almenningur að borga brúsann þegar fyrirtæki fara á hausinn; borga (e-ð[upphæð, mikið/lítið]) fyrir e-ð[hlut, vöru, þjónustu]; borga e-ð[lán, skuld] upp þú losnar við frekari vexti ef þú borgar lánið upp strax; borga e-ð[upphæð] út (í e-u[hlut, vöru, fasteign]) hann bauðst til að borga út fjórðung í blínum → útborgun 3. <e-að[framkvæmd, verknaður]> borga sig (ekki) það (marg)borgar sig að láta gera við sjónvarpið, það borgar sig ekki að svara honum þegar hann er að skammast. 4. borga e-m/e-u[fyrirtæki, stofnun] (e-ð) (fyrir e-ð) ég borgaði smíðnum 1000 krónur á tímann; borga e-m (e-ð[upphæð]) út → útborgun*
- "m. akk."  
"ordforb."
- "m. dat.  
(+ akk.)"
- borgun

"m. akk."

**skapa** so <e-τ/(e-að)> 1. skapa e-n[ílfveru, manninn ...]/e-ð[heiminn; listaverk; aðstæður, ástand, vanda, vandamál; eftirspurn ...] *guð skapaði manninn í sinni mynd, mótmæli hvalfríðunarmanna hafa skapað mikinn vanda meðal þeirra sem byggt hafa afkomu sína á hvalveidum, harkaleg afstaða stjórnvalda hefur skapað mikla óánægju í röðum launamanna* 2. <e-τ/e-að[stjórnvöld, yfirvöld; aðgerðir ...]> skapa e-m/e-u[starfsemi] e-ð[aðstöðu, skilyrði] *endurbætur á höfninni hafa skapað sjómönnum allt aðrar og betri aðstæður en áður, með gengisbreytingunni hafa úflutningsatvinnuvegunum verið sköpuð viðunandi skilyrði; skapa sér e-ð[aðstöðu, skilyrði, sérstöðu ...] sem ritari ráðherrans skapaði hún sér ákjósanlega aðstöðu til njósna.*

"m. dat.  
+ akk."

-skapa: á-, endur-, um- ...

"ordforb"

skapast miðm e-að[ástand, aðstæður, vandamál ...] skapast (með e-u/við e-ð) *eftir byilinguna sköpuðust aðstæður til að auka lýðræði í landinu.*

skapaður lh þt (ekki) nokkur skapaður hlutur hann gat engu svarað, vissi ekki nokkurn skapaðan hlut, hún hefur ekki áhuga á nokkrum sköpuðum hlut; ekki nokkurn skapaðan hlut þér kemur þetta ekki nokkurn skapaðan hlut við; allir skapaðir hlutir hann er ákaflega forvitinn, spyr um alla skapaða hluti, hún hefur áhuga á öllum sköpuðum hlutum.

-skapaður: 1. á-, sjálf-, van- ...2. al-, fagur-, full-, hálf-, ftur- ...

skapandi lh nt 1. <e-τ[listamaður/höfundur ...]> 2. <e-að[list, hugsun, fmynd-unarafi, starf ...]>.

→ sköpun, skapari

### III. Adjektiv

**fagur** lo 1. <e-τ> vera fagur (ásýndum) *furstinn naut þess að vera umkringdur fjögurum konum* 2. <e-að[listaverk ...; sjón, útsýni; staður, landslag; veður; dagur; ræða, lýsing, orð ...]> skáldinu var færður fagur silfurskjöldur í viðurkenningar-skygni, það er fagurt útsýni af jöklinum, afmælisbarnið þakkaði fögur orð í sinn garð, við ókum af stað í fegursta veðri 3. <e-að[ástand; lýsing, orð, ummæli ...]> vera ekki fagurt margt af því sem hann sagði var ekki fagurt 4. það er fagurt e-τs staðar það er fagurt hér í eyjunum í björtu veðri.

"forsterkende"

-fagur: -blár, -bókmenntir, -eyg(ð)ur, -grænn, -gulur, -lega, -orður, -rauður ...

-fagur: dá-, engil-, forkunnar-, fól-, stór-, unaðs-, undra-, undur-, yndis-, ægi-

...

→ fegurð

**saklaus** lo 1. <e-τ> *viðurkennir hann þjófnadinn? - nei, hann segist vera saklaus; vera saklaus af e-u ég er saklaus af því að hafa misnotað aðstöðu mína* 2. <e-að[gerð, verknaður]> *er ekki saklaust að fá sér einn bjór?, strákarnir lítu á slagsmál sem saklausa skemmtun; e-m er saklaust að gera e-ð.*

"forsterkende"

-saklaus: al-, sára-, vita-

→ sakleysi

&lt;◇&gt; sekur

**sekur** lo <e-τ> (vera) sekur (um e-ð[afbrot, glæp]); gera sig sekan um e-ð[yfirsjón, afbrot]; vera sekur í e-u[glæp]; dæma e-n sekan.

-sekur: dauða-, marg-, með-, stór- ...

→ sekt

&lt;◇&gt; saklaus, sýkn

Grete Duvå, Anna-Lise Laursen, Lisbet Maidahl

## Brugerundersøgelse vedrørende oversættelse af fagtekst

Som første fase i et leksikografisk projekt, der gennemføres på Handelshøjskolen i Århus, er der foretaget en undersøgelse, der fokuserer på brugen af ordbøger. Undersøgelsens formål har været at kortlægge de problemer, som opstår under oversættelse af fagtekster, at få viden om, hvordan oversætteren løser problemerne, og hvilke hjælpemidler vedkommende benytter sig af. Informanterne er sprogstuderende, økonomistuderende og professionelle oversættere fordelt på sprogene fransk, spansk og tysk. Informanterne har i protokolform beskrevet problemer og løsninger, og på basis af deres oplysninger er der foretaget en analyse af brugerstrategierne og en typologisering af problemerne. Resultatet af analysen kan betragtes som et forslag til en forbedring af den eksisterende tosproglige fagleksikografi.

### 1. del:

Lisbet Maidahl

#### Analyse af brugerundersøgelsen med tysk som målsprog

##### 1 Baggrund for undersøgelsen

Fra 1990 til 1993 gennemføres i Danmark projektet "Oversættelse af fagsproglige tekster" som et samarbejde mellem handelshøjskolerne i København og Århus, universiteterne i København og Århus, Danmarks Lærerhøjskole og Eurotra, finansieret af det Humanistiske Forskningsråd. Inden for rammerne af dette projekt beskæftiger vi os på Handelshøjskolen i Århus med oversættelsesorienteret leksikografi og har i den forbindelse gennemført en brugerundersøgelse vedrørende oversættelse af fagtekster for at få et nøjere kendskab til problemerne under oversættelsesprocessen, fremgangsmåden ved problemløsningen og hjælpemidlernes art og ydeevne i denne sammenhæng.

Projektet på Handelshøjskolen i Århus falder naturligt ind i den udvikling, der i løbet af de seneste år er sket på det leksikografiske område, hvor der sideløbende med den hidtidige interesse for almensproglige ordbøger har vist sig en begyndende interesse for fagsproglig leksikografi, hvilket har bevirket, at der er ved at udvikle sig en gren inden for den traditionelle leksikografi, hvis tyngdepunkt er fagsproget. Da fagsprog stiller andre krav til ordbøger end almensproget, må denne særlige fagleksikografi formodes at udvikle teorier, der tilgodeser disse specifikke krav, som naturligvis bør afspejle ordbogsbrugerens behov.

Men hvori består disse behov? Udgangspunktet for bilingual fagsproglig leksikografisk forskning må være en viden om fagoversætterens problemer og behov, og på denne baggrund tager vores brugerundersøgelse direkte sigte på at belyse oversætterens brug af hjælpemidler.

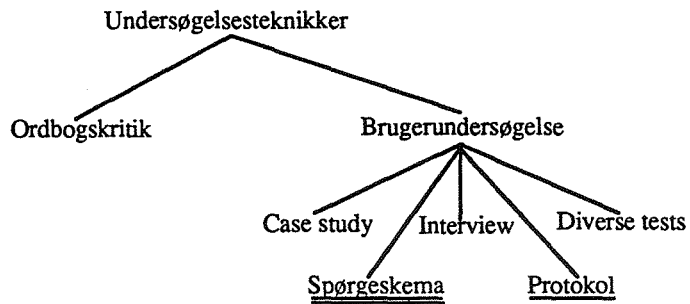
Det overordnede formål med undersøgelsen er på basis af det empiriske materiale at foretage en analyse af fagoversætterens søgestrategier og problemregistreringer og anvende resultaterne heraf til udarbejdelse af et koncept til principperne for en ideel fagordbog inden for økonomisk sprog idet undersøgelsens fagområde er økonomi inden for sprogene fransk, spansk og tysk.

Under projektet er der desuden gennemført brugerundersøgelser for fagområderne teknik og jura inden for sprogene spansk, engelsk og tysk med en informantgruppetørrelse på ca. 10 til 20 personer.

## 2 Brugerundersøgelsen

Før gennemførelsen af undersøgelsen har en afvejning af forskellige undersøgelsesteknikker i relation til undersøgelsesformålet været nødvendig.

Nedenstående oversigt over forskellige undersøgelsesteknikker skal illustrere mulige fremgangsmåder.



Figur 1

I størstedelen af undersøgelser af ordbogsbrugere arbejdes der med spørgeskemaet, eventuelt i kombination med andre metoder. Af de senere undersøgelser, der har valgt spørgeskemaet, kan henvises til Ripfel (1989), Benbow et al. (1990) eller Bogaards (1990), mens kun få, bl.a. Wiegand (1985) og Neubauer (1987), har baseret deres undersøgelse på protokolmetoden. Ripfel/Wiegand (1988) giver en kommenteret oversigt over de hidtil gennemførte undersøgelser, på samme måde Diab (1990), hvor oversigten suppleres med en kritisk vurdering af de mulige metoder.

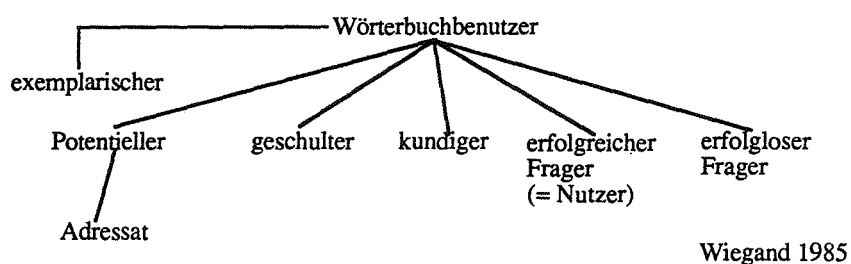
Vi har valgt protokolmetoden

- 1) for at undgå at styre besvarelsenerne
- 2) for at komme så tæt som muligt på en "naturlig" oversættelsessituation og
- 3) for at få beskrevet flest mulige af de problemer, oversætteren/ordbogsbrugeren selv identificerer

Denne er suppleret med interviewmetoden, for så vidt angår den gruppe informanter, der oversatte til fransk. Protokollen udfyldes i direkte forbindelse med, altså sideløbende med

oversættelse af en fagtekst. For at opnå den mest virkelighedstro oversættersituation er oversættelsen og udfyldningen af protokollen foregået på det for informanten mest naturlige sted, altså hjemme, på arbejde, på bibliotek eller i undervisningen. Som fagtekst til oversættelse er valgt et uddrag af Bang & Olufsens årsregnskab 1989 (jf. bilag 1 og side 000), dvs. en teksttype, som professionelle hyppigt arbejder med.

Informantgruppen har således bestået af sådanne personer, der gennem deres daglige arbejde eller under uddannelsen beskæftiger sig med de fagområder, teksten berører. Det drejer sig om professionelle oversættere, studerende fra oversætter- og tolkeuddannelsen og økonomistuderende med et godt kendskab til de respektive sprog. Dvs. alle vores informanter har et sprogniveau, som Wiegand (1985) betegner som "der geschulte Wörterbuchbenutzer":



Figur 2

I tabel 1 vises et eksempel på en udfyldt protokol med tilhørende liste over anvendte hjælpemidler. Protokollen er delt i 4 kolonner. I første kolonne noteres det nummer, informanten har givet problemet i fagteksten, dernæst beskrives eller nævnes problemets art. I anden kolonne anføres i form af kode, hvilke hjælpemidler der blev brugt til løsning af samme problem, idet hjælpemidlerne noteres i den rækkefølge, de er anvendt i. Desuden afskrives en del af oplysningen, der blev brugt. I tredje kolonne kommenteres den fundne oplysnings brugbarhed og udførlighed med henblik på grammatiske oplysninger, definitioner, eksempler o.l. Til sidst angives i fjerde kolonne den til løsning af problemet medgåede tid.

KOMMENTARER TIL OVERSÆTTelsesPROBLEMER					
KOLONNE 1		KOLONNE 2		KOLONNE 3	KOLONNE 4
OVERSÆTTelsesPROBLEM		HJÆLPEMIDLER (incl. mundtlig hjælp)		BEMÆRKNING TIL DEN FUNDNE OPLYSNING	Det tog ca. x min. at løse problemet
nr.	Beskrivelse	Hvilken oplysning gav hjælpemidlerne	kode		
10	udbytter	- konzerninterne Gewinne - - -	C D		8
11	interne aktie- besiddelser	- jeg slår op under noget med "Aktien" - intet resultat - - -	A C D		6
12	mellemvæ- render	- under "Jahresabschluss" finder jeg "Beteiligungen", som jeg går videre med i D. - her findes "konzerninterne Beteiligungen" - Forderungen und Verbindlichkeiten	E D A	checkes i D	12
13	konzerndæk- ningsbidrag	- her optræder "konzerninterne Forderungen und Verbindlichkeiten" - da jeg ved, "dækningsbidrag" hedder "Deckungsbeitrag", søger jeg under denne term, men finder det desværre ikke - heller intet resultat - - - da jeg heller ikke her finder den ønskede sammensætning, har jeg 3 muligheder: 1) lave sammensætningen "Konzerndeckungsbeitrag" ud fra den betragtning, at der findes mange sådanne sammensætninger i kilderne med Konzern- 2) lave: Deckungsbeitrag des Konzerns, da der f.eks. bruges "die Lage des Konzerns" Da der findes så mange komposita med "Konzern-", og det desuden ikke kan misforstås, vælger jeg løsning 1.	D A C		15

Tabel 1

Ved besvarelsen var alle hjælpemidler tilladt, og som det fremgår af tabel 2 har denne informant anvendt syv forskellige, fordelt på ordbøger, lærebøger og leksika; desuden har der af andre informanter også været anvendt paralleltekster, brochurer o.a.



LISTE OVER ANVENDTE HJÆLPEMIDLER (BIBLIOGRAFI)					
kode	FORFATTER/ORG. el. lign.	TITTEL	Udgiv. år	ART (ordbog, lærebog, brochure etc.)	Din vurdering af kilden (god, dårlig, ikke ud- førlig nok el. lign.)
A	W, Gubbas Forlag, Mårslet	juridisk ordbog	1984	ordbog	uundværlig
B	Dudenverlag	Duden - das grosse Wörter- buch der deutschen Sprache	1977	ordbog	uundværlig, men gerne flere ex.
C	Verlag Vahlen	Einführung in die Allgemeine Betriebswirtschaftslehre	1986	lærebog	omfattende og derfor tidskræ- vende
D	Gabler	Konzernabschlüsse	1984	lærebog	omfattende og tidskrævende
E	Betriebswirtschaftlicher Verlag	Gablers Wirtschaftslexikon	10. Aufl.	lexikon	meget, meget god
F	A.M. Kjærs Forlag	Handelsordbog (Poulsens)	1983	ordbog	god, pålidelig, gerne flere ter- mer
G	Gjellerup & Gad	Virksomhedsøkonomi	1987	lærebog	udmærket af en dansk lærebog at være

Tabel 2

### 3 Komparativ analyse af brugerundersøgelsen med tysk som målsprog

#### 3.1 Kvantitativ analyse

Et uddrag af Bang & Olufsens årsregnskab 1989 med tilhørende udfyldte skemaer til protokol og "liste over hjælpemidler" (jf. tabel 1 og 2) blev givet til to forskellige informantgrupper. Den ene er sprogstuderende i 6. og 8. semester, der har aflagt eksamen i dansk regnskabslære i 2. semester, dvs. en gruppe med et forholdsvis højt sprogniveau og et mindre fagkendskab, her kaldet sproginformanter. Den anden informantgruppe, her kaldet økonomiinformanter, består af økonomistuderende på 6., 7. og 8. semester, der har haft hhv. 10, 14 og 16 semestertimer tyskundervisning og har afsluttet regnskabsvæsen i 4./5. semester. Der er tale om en informantgruppe med et højere fagkendskab og et lavere sprogniveau end sproginformanterne. Der er modtaget besvarelser fra 9 sproginformanter og 18 økonomiinformanter. Nedenstående opstilling viser en kvantitativ oversigt over det totale antal problemer, det totale antal opslag til løsning af problemerne samt den totalt medgåede tid.

## PROBLEM

		Problemer/ opslagsord totalt	Antal opslag	Tidsforbrug totalt til løsning af problemet
SPROG i alt (9)	CLM <sup>1</sup>	126	177	426 min.
ØKONOMI i alt (18)		358	448	587 min.
	CMI (6) <sup>2</sup> (7.+8. sem.)	106	135	191 min.
	HE (12) <sup>3</sup> (6. sem.)	252	313	397 min.

1) kandidatuddannelse erhvervsprog (100% sprog)

2) kandidatuddannelse økonomi og sprog (60% økonomi, 40% sprog)

3) basisuddannelse økonomi og sprog (60% økonomi, 40% sprog)

Tabel 3

Af tabel 3 fremgår det, at 18 økonomiinformanter i forhold til 9 sproginformanter har ca. tre gange så mange problemer og opslag til løsning af problemerne, mens tidsforbruget kun er en fjerdedel større. Til tydeliggørelse af disse forhold reduceres økonomiinformanternes tal til det halve, hvorved grundlaget bliver et lige stort antal informanter.

## EN SAMMENSTILLING AF LIGE STORE INFORMANTGRUPPER

		Problemer/ opslagsord totalt	Antal opslag	Tidsforbrug totalt til løsning af problemet
SPROG i alt (9)	CLM	126	177	426 min.
ØKONOMI i alt (9)		179	224	293,5 min.

Tabel 4

En sammenligning af tabellerne aftegner en vis tendens i forbindelse med problemer/opslagsord og løsning af dem, som vil blive illustreret i den følgende beskrivelse og fortolkning af det kvantitative talmateriale.

*Problemer/opslagsord:* Der er her tale om det samlede antal ord, de to informantgrupper har beskrevet som et problem. Sammenligner man de 9 sproginformanternes 126 problemer/opslagsord med økonomiinformanternes 179, ser man, at den sidstnævnte gruppe har 53 flere beskrevne problemer/opslagsord eller ca. 50% flere.

*Antal opslag:* Denne kolonne angiver det samlede antal opslag, der er foretaget for at løse de beskrevne problemer, altså finde en rigtig oversættelse. Sproginformanterne bruger 177 opslag til at finde en løsning. Økonomiinformanterne derimod 224 opslag, altså 43 flere.

Dette betyder, at selv om økonomiinformanterne iflg. 1. kolonne ved oversættelsen til tysk har ca. 50% flere problemer, bruger de kun en fjerdedel flere opslag til at løse dem. Med andre ord, informanter med bedre fagkundskab synes hurtigere at finde en løsning på deres problem.

*Tidsforbrug:* I denne kolonne er opført det samlede antal minutter, der iflg. informanterne er medgået til at finde oversættelserne til opslagsordene. Man kan heraf konkludere, at den på baggrund af 2. kolonne konstaterede tendens i retning af hurtigere løsninger nu bliver yderligere belyst med minutangivelser. De 9 økonomiinformanter bruger nemlig godt 133 min., eller godt 30% mindre tid, til at løse ca. 50% flere problemer/opslagsord. Man må på baggrund heraf spørge, om resultatet af de to gruppers løsningsforslag også er betydeligt forskelligt.

Opstillingen i tabel 5 er baseret på både informantgruppernes oplysning om, hvilken løsning de valgte, og på deres afleverede oversættelse. De to første kolonner viser det samlede antal rigtigt løste problemer og det samlede antal ikke løste, i tredje kolonne er forholdene udtrykt i procent. Tallene er baseret på 9 sproginformanter og 18 økonomiinformanter.

#### Løsningsforslag:

##### Registrerede opslagsord

		løst	ikke løst	løst i %
SPROG	CLM	99	22	78
i alt (9)				
ØKONOMI		232	126	65
i alt (18)				
	CMI (6) (7.+8. sem.)	75	31	
	HE (12) (6. sem.)	157	95	
				diff. 13%

Tabel 5

Som det fremgår af opstillingen, løser sproginformanterne ca. 78% af deres problemer/opslagsord, økonomiinformanterne derimod 65%. Der er altså tale om en difference på 13% flere korrekte løsninger for sproginformanternes vedkommende. I tabel 6 vises en oversigt over, hvad det gennemsnitlige tidsforbrug til det gennemsnitlige antal opslag har været for de to informantgrupper ved oversættelsen af fagteksten.

	Sproguddannede CLM (9)	Økonomiuddannede (CMI (6)	HE (12)
<b>Antal opslag</b>			
i alt	177	135	313
i snit pr. problem	1,4	1,21	1,22
<b>Tidsforbrug pr. problem</b>	5 min.	1,9 min.	2,6 min.
i snit			
min./max.	1-42 min.	0,5-10 min.	0,5-45 min.
<b>Tidsforbrug totalt</b>			
i snit	4 timer	4,5 timer	4,8 timer
min./max.	1,5-10 timer	0,4-8 timer	2,4-10 timer

- sproginformanterne foretager i snit 1,4 opslag, hvortil de bruger i snit 5 min.
- økonomiinformanterne foretager i snit 1,1, opslag, hvortil de bruger 2,3 min. i snit

Tabel 6

Tabel 6 viser at sproginformanterne foretager i snit 1,4 opslag, hvortil de bruger i snit 5 min. og at økonomiinformanterne foretager i snit 1,1 opslag, hvortil de bruger ca. 2,3 min. i snit. Sproginformanterne bruger altså dobbelt så lang tid, selv om de har et noget højere sprogniveau, til at finde 13% flere rigtige løsninger på deres problem end informantgruppen med den større faglige indsigt. Jeg mener, at man heraf kan slutte, at en sproginformant med et højere sprogniveau end økonomiinformanten i sine ordbøger ikke får de definitioner og encyklopædiske oplysninger, der sætter hende/ham i stand til at oversætte en typisk tekst i et oversættelsesbureau korrekt på en rimelig tid. Der må være behov for langt flere og mere entydige informationer i ordbøgerne.

### 3.2 Kvalitativ analyse

#### 3.2.1 Problemer i fagteksten

På baggrund af det forudgående er det måske overraskende, at informantgrupperne har angivet næsten de samme ord og ordforbindelser som værende et problem. I tabel 7 som viser afskrift af fagteksten er de registrerede ord eller ordforbindelser nummereret fortløbende, idet den kantede parentes omslutter ordforbindelsen.

Opslagsordene/problemerne er næsten identiske, dog optræder 5 opslagsord mere i økonomigruppen, nemlig *omfatter*<sup>2</sup>; *for såvel ... som for*<sup>4</sup>; *udarbejdet*<sup>4</sup>; *i de udenlandske*<sup>21</sup>; *pr. 31. maj*<sup>31</sup>. Kun ved nr. 14 *udarbejde* kan man tale om et problem, der beror på tekstens fagområde, koncernregnskaber, nemlig et kollokationsproblem ved syntagmet, *årsregnskabet/koncernregnskabet er udarbejdet*, her vælges *aufstellen* på tysk. De øvrige fire problemer er af almensproglig art og vil indgå i en senere analyse.

SPROGINFORMANTER (CLM)	ØKONOMIINFORMANTER (HE/CMI)
<p><b>Anvendt regnskabspraksis<sup>1</sup></b></p> <p>Årsregnskabet<sup>2</sup> omfatter regnskab<sup>3</sup> for såvel koncernen<sup>4</sup> som for Bang &amp; Olufsen A/S, og er udarbejdet<sup>5</sup> [i overensstemmelse med]<sup>6</sup> årsregnskabsloven<sup>7</sup> af 10. juni 1981.</p> <p>Regnskabet er udarbejdet efter samme regnskabspraksis<sup>9</sup> som [året<sup>8</sup> før].</p> <p>Konsolidering:<sup>10</sup></p> <p>Koncernregnskabet<sup>11</sup> er udarbejdet som et [sammendrag af]<sup>12</sup> de enkelte selskabers<sup>13</sup> regnskabsposter.<sup>14</sup> Der er foretaget eliminering<sup>15</sup> af [intern omsætning],<sup>16</sup> interne<sup>17</sup> renter og udbytter,<sup>18</sup> interne aktiebesiddelser<sup>19</sup> og mellemværender<sup>20</sup> samt tilbageførsel<sup>21</sup> af ikke realiseret<sup>23</sup> koncerndækningsbidrag.<sup>22</sup> Regnskabsposterne<sup>24</sup> i de udenlandske datterselskaber er omregnet<sup>25</sup> efter valutakurser<sup>26</sup> pr. 31. maj.</p>	<p><b>Anvendt regnskabspraksis<sup>1</sup></b></p> <p>Årsregnskabet<sup>2</sup> omfatter<sup>3</sup> regnskab<sup>5</sup> [for<sup>4</sup> såvel] koncernen<sup>6</sup> [som for]<sup>4</sup> Bang &amp; Olufsen A/S, og er udarbejdet<sup>7</sup> [i overensstemmelse med]<sup>8</sup> årsregnskabsloven<sup>9</sup> af 10. juni 1981.</p> <p>Regnskabet er udarbejdet efter samme regnskabspraksis<sup>10</sup> [som året før].<sup>11</sup></p> <p>Konsolidering:<sup>12</sup></p> <p>Koncernregnskabet<sup>13</sup> er udarbejdet<sup>14</sup> som et [sammendrag af]<sup>15</sup> [de enkelte selskabers]<sup>16</sup> regnskabsposter.<sup>17</sup> Der er [foretaget eliminering]<sup>18</sup> af]<sup>18</sup> [intern omsætning],<sup>19</sup> interne renter<sup>20</sup> og udbytter,<sup>21</sup> interne aktiebesiddelser<sup>22</sup> og mellemværender<sup>23</sup> samt [tilbageførsel af]<sup>25</sup> ikke realiseret<sup>26</sup> koncerndækningsbidrag.<sup>24</sup> Regnskabsposterne [i de udenlandske]<sup>27</sup> datterselskaber<sup>28</sup> er [omregnet efter]<sup>29</sup> valutakurser<sup>30</sup> [pr. 31. maj.]<sup>31</sup></p>

Tabel 7

### 3.2.2 Problemart

Ser man i tabel 8 på de problemer/opslagsord, informanterne har foretaget, og læser man, hvordan de selv i deres protokol har beskrevet deres problemer under løsningsforsøget, fremkommer først en fordeling på problemarterne ækvivalens og grammatik. Ved ækvivalensproblem menes, at informanten spørger "hvad hedder ...?", som grammatisk problem opføres genus, deklination, konjugation, rektion. Betegnelserne dækker i første omgang informanternes egne beskrivelser.

Problem:

		Opslagsord totalt	Problemart	
SPROG i alt (9)	CLM	126	147	142 Æ <sup>1</sup> 5 G <sup>2</sup>
ØKONOMI i alt (18)		358	503	441 Æ 62 G
	CMI (6) (7.+8. sem.)	106	150	127 Æ 23 G
	HE (12) (6. sem.)	252	353	314 Æ 39 G

1) ækvivalensproblem

2) grammatisk problem

Tabel 8

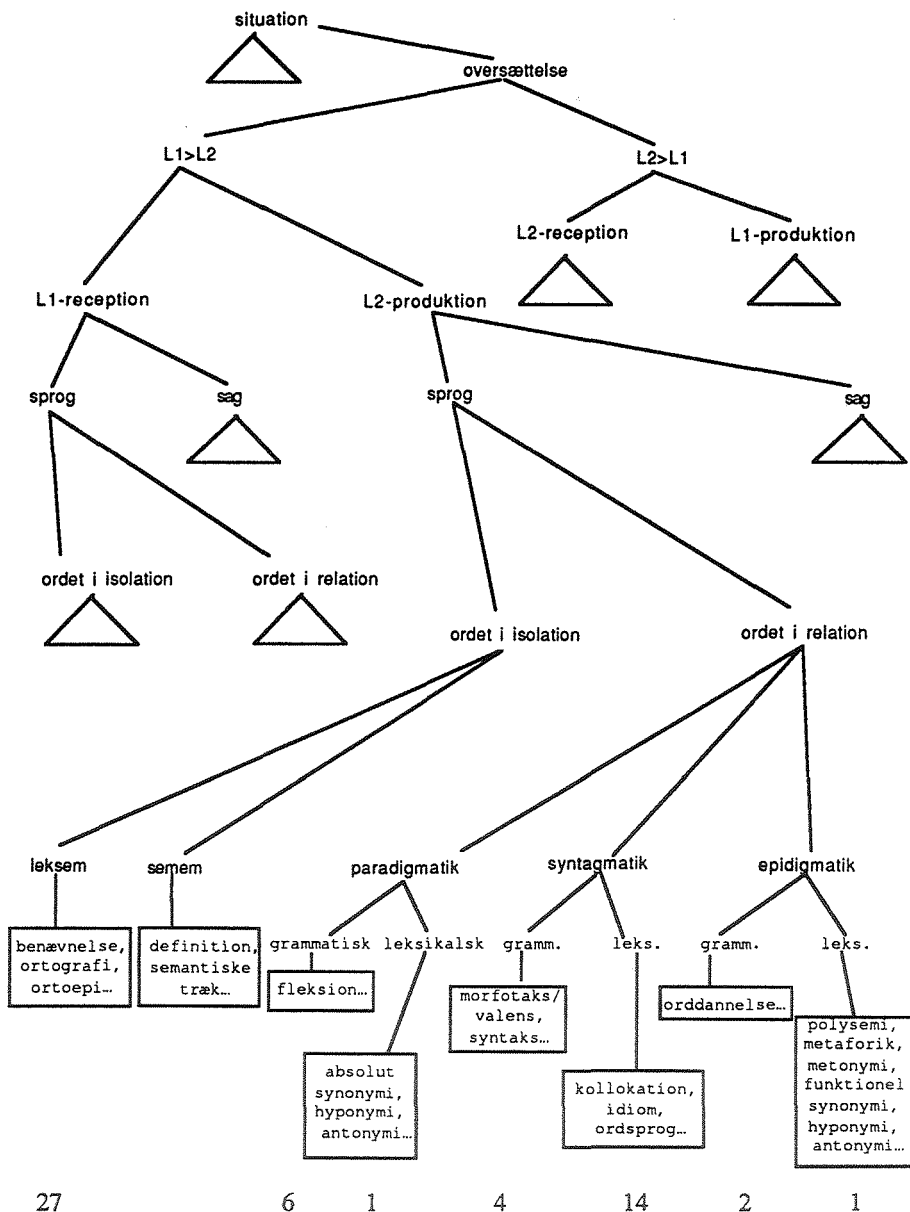
Det fremgår heraf, at sproginformanterne efter deres oplysninger har ganske få grammatiske problemer, nemlig 5, så som substantivets genus ved "koncern", mens økonomiinformanterne i alt beskriver 62 grammatiske problemer. Det noget større antal problemarter i forhold til opslagsord skyldes, at der ved et problem/opslagsord ofte har været flere problemarter.

En første konklusion på basis af den kvalitative analyse er, at der i begge informant-grupper fortrinsvis hersker usikkerhed over for brugen af den rette fagterminologi.

### 3.2.3 Problemløsningsstrategi og problemtypologi

Da den forudgående inddeling i ækvivalens- og grammatikproblemer, beroende på informanternes egne beskrivelser, er ret grov, forsøges i det følgende foretaget en mere præcis fastlæggelse af problemløsningsstrategi og problemtypologi på baggrund af informanternes opslag. Til denne analyse er anvendt en interpretationsmodel, som er præsenteret i figur 3. Modellen er opstillet i samarbejde med mag.art. H.K. Mikkelsen, som deltager i det leksikografiske projekt på Handelshøjskolen i Århus.

Hans Kristian Mikkelsen: PROBLEMTYPOLOGI - EN INTERPRETATIONSMODEL



Figur 3

I modellen står  $L_1$  for udgangssprog,  $L_2$  for målsprog. De til sidst i interpretationsmodellen anførte tal angiver, hvor ofte denne problemtype/problemløsningsstrategi er forekommet i forbindelse med de 31 problemer i fagteksten, som sprog- og økonomiinformanterne har noteret.

Ved opdelingen i  $L_2$  ord i isolation og  $L_2$  ord i relation får man to kategorier af ca. samme størrelse, nemlig hhv. 27 og 28.  $L_2$  i isolation som benævnelse, altså fagtermer på  $L_2$  sproget, er for begge informantgrupper hovedproblemet, hvilket understreger betydningen af præcise forklaringer, definitioner og encyklopædiske oplysninger i fagordbøger. Samtidig viser kategorien  $L_2$  ord i  $L_2$  relationer, fordelt på både leksikalske og grammatiske relationer, at den relation,  $L_2$  ordet hovedsagelig søges i, er leksikalske syntagmer i form af kollokationer (14). Den under leksikalsk paradigmatik opførte problemtype er et synonymproblem på  $L_2$ .

Under relationen grammatisk epidigmatik søges der i disse eksempler dannet komposita, f.eks. med "Konzern" og "Deckungsbeitrag". De øvrige grammatiske problemtyper fordeler sig på grammatisk paradigmatik (6), her fleksion, og grammatisk syntagmatik (4), her dannelse af syntagmer som syntaktisk størrelse på  $L_2$ . Der er her ikke optegnet, hvordan disse problemtyper/problemløsningsstrategier fordeler sig på de 27 informantere i alt 583 ækvivalensproblemer og i alt 67 grammatiske problemer, men kun hvor mange forskellige problemtyper de 31 problemer i teksten udgør.

Alle problemer/opslagsord er forsøgt interpreteret med udgangspunkt i informantens udfyldning af kolonne 1, 2 og 3 i skemaet "kommentarer til oversættelsesproblemer" (jf tabel 1) i indledningen). De nedenfor opførte 5 typeeksempler illustrerer interpretationer af sprog- og økonomiinformanternes problemer/opslagsord fra fagteksten, uden sammenligning; nummereringen henviser til økonomiinformanternes nummerering (jf. tabel 7).

### Typeeksempler på oversættelsesproblemer

#### *14 sammendrag/sammendrag af*

Opslagsord: 17  
 Antal opslag: 26  
 Løst: 16

Problemtypologi/problemløsnings-strategi:

- 1)  $L_2$  ord i isolation som term ved koncerntregnskab
- 2)  $L_2$  ord i  $L_2$  relation: grammatisk syntagmatik: "af" til dannelse af substantivisk syntagme, på  $L_1$ , der søges den tilsvarende syntaktiske størrelse på  $L_2$ .

Kommentar:

Findes ved opslag i 2-sproget fagordbog d/t Gubba med eksempel, men mange læser lærebøger og leksika, især sproginformanterne.

Løsning: eine Zusammenfassung der ... ist aufgestellt.



*15 de enkelte selskaber/selskaber*

Opslagsord: 7  
 Antal opslag: 12  
 Løst: 6

## Problemtypologi/strategi:

L<sub>2</sub> ord i isolation som term ved koncerner  
 L<sub>2</sub> ord i isolation som almensprogligt ord

## Kommentar:

Ved opslag i 2-sproget fagordbog d/t Gubba findes "Konzernunternehmen" som term ved koncern, ikke "Gesellschaft".

Løsning: die einzelnen Konzernunternehmen

*18 elimineringsforetagelse elimineringsforetagelse elimineringsforetagelse*

Opslagsord: 15  
 Antal opslag: 25  
 Løst: 14

## Problemtypologi/problemløsnings-strategi:

- 1) L<sub>2</sub> ord i isolation som term ved koncernregnskab
- 2) L<sub>2</sub> ord i L<sub>2</sub> relation: leksikalsk syntagmatik "foretagelse" som L<sub>2</sub> kollokation ved "Verrechnung"
- 3) L<sub>2</sub> ord i L<sub>2</sub> relation: grammatisk epidigmatik som substantivering af "eliminieren".
- 4) grammatisk syntagmatik: "af" til dannelse af substantiviske syntagmer, der søges den tilsvarende syntaktiske størrelse på L<sub>2</sub>.

## Kommentar:

2-sproget fagordbog d/t Gubba angiver verbet "eliminieren", t/d Gubba derimod "Verrechnung". "Eliminierung" dannes i analogi til "Verrechnung" af flere, da det ikke er et opslagsord i fagordbogen.

Løsning: die Verrechnung/die Eliminierung

*19 intern omsætning*

Opslagsord: 17  
 Antal opslag: 21  
 Løst: 14

## Problemtypologi/problemløsnings-strategi:

L<sub>2</sub> ord i L<sub>2</sub> relation: leksikalsk syntagmatik på L<sub>1</sub>, der på L<sub>2</sub> optræder som grammatisk epidigmatik = kompositum intern omsætning = Innenumsatz (m.)  
 → term ved koncernregnskab

## Kommentar:

Problemet kan opfattes som en term ved koncernregnskab på tysk og kan slås op i 2-sproglig fagordbog d/t-t/d E. Fryd

Løsning: der Innenumsatz

*28 valutakurs*

Opslagsord: 11  
 Antal opslag: 13  
 Løst: 2

## Problemtypologi/problemløsnings-strategi:

L<sub>2</sub> ord i isolation som term ved valuta.

## Kommentar:

Som opslagsord i 2-sprogede fagordbøger findes ikke den rigtige løsning, mangler eksempler.

Løsning: der Wechselkurs

## Citerede hjælpemidler:

Gubba, Wilhelm (1984), *Juridisk ordbog, dansk-tysk*

Fryd, Ejnar (1974), *Special-ordbog inden for revision, regnskabsvæsen m.v., dansk- tysk, tysk dansk*

På baggrund af denne analyse må man ønske, at tosprogede fagordbøger i højere grad kommer til at dække oversætterens behov. I den kvantitative analyse blev der på basis af en sammenligning af sproginformanter og økonomiinformanter konkluderet, at den medgåede tid ved søgningen af den rigtige fagsproglige term må kunne begrænses ved flere definitioner, forklaringer og encyklopædiske oplysninger i fagordbøger. Dette understreges af den kvalitative analyse, der også viser, at fagterminologien er hovedproblemet. Men for at tilgodese det ligeledes tydelige behov inden for problemtypen leksikalsk syntagmatik er det desuden nødvendigt, at fagordbogen indeholder gode eksempelsætninger, der illustrer fagtermens leksikalske L2-relationer. Derudover blev det klart, at fagtermen også bør

beskrives og illustreres i alle grammatiske L2-relationer. En fagordbog, der tilgodeså disse behov, ville være til stor hjælp for alle, der i deres profession eller under uddannelsen beskæftiger sig med vanskelige fagtekster.

## 2. del:

Grete Duvå & Anna-Lise Laursen

### Analyse af brugerstrategier og problemarter med fransk og spansk som målsprog

#### 1 Informanter

##### 1.1 Informantkategorier

Som informanter er anvendt studerende, færdige kandidater, der er sprogmedarbejdere i virksomheder, samt professionelle oversættere. Disse kategorier formodes så nogenlunde at svare til de potentielle brugere af en eventuel fagordbog. Informantgrupperne er sammensat som vist i tabel 1:

Sprog	Kategori	Antal
Fransk	CLM 1. sem	7
Fransk	CLM 4. sem	7
Fransk	færdige kand/prof oversættere	6
<b>Fransk</b>	<b>I alt</b>	<b>20</b>
Spansk	CLM 1. sem	7
Spansk	færdige kand/prof oversættere	7
<b>Spansk</b>	<b>I alt</b>	<b>14</b>

Tabel 1

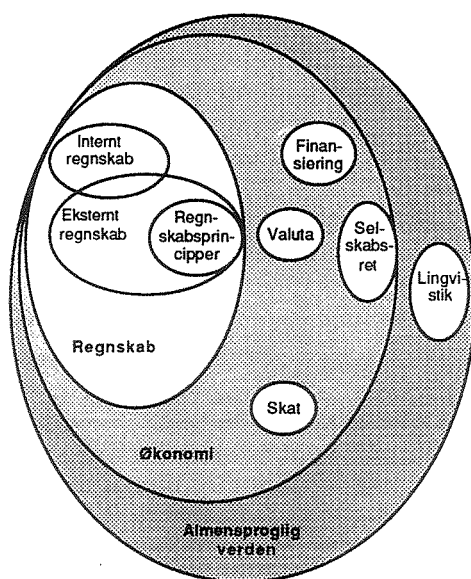
CLM (cand.ling.merc.) er en 5 1/2 årig oversætter- og tolkeuddannelse, der består af en 1. del (korrespondentuddannelse) på 2 1/2 år og en overbygning på 3 år. Informanterne befinder sig på hhv. 1. og 4. semester på overbygningen. For begge sprog var situationen i øvrigt den, at ingen af grupperne (bortset fra de færdiguddannede) havde særlige forudsætninger i fransk og spansk til at løse den stillede opgave, dvs. de havde endnu ikke beskæftiget sig med fagområdet i fransk-, hhv. spanskundervisningen. De formodes derimod alle at have haft introducerende undervisning i emnet på erhvervsøkonomi på korrespondenttrinnet.

## 2 Fagtekst

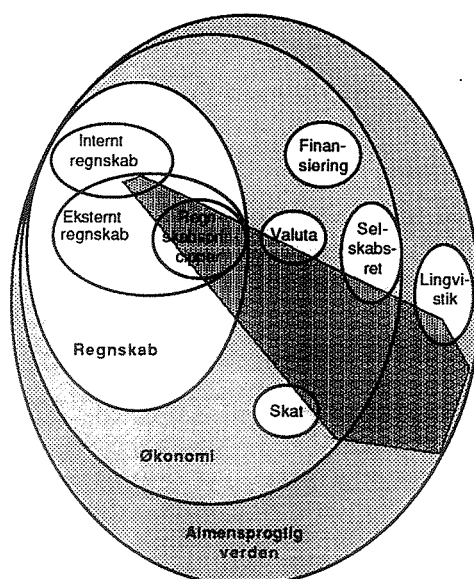
Den tekst, informanterne fik til opgave at oversætte, hører til det økonomiske fagområde, nærmere betegnet eksternt regnskabsvæsen (teksten vises på bilag 1). Den er fra Bang & Olufsens årsberetning, og emneområdet er virksomhedens regnskabsprincipper. Teksten repræsenterer en typisk økonomisk tekst, da bl.a. virksomheders årsberetninger alle indeholder afsnit om, efter hvilke regnskabsprincipper virksomheden har opstillet sit regnskab.

Da virksomhedernes årsberetninger desuden spiller en stor rolle på det internationale kapitalmarked bl.a. som vurderingsgrundlag for internationale investorer, er det en tekst som typisk vil blive oversat i virksomhederne eller af oversættelsesbureauer. For at kunne løse oversættelsesopgaven tilfredsstillende hævder vi på grundlag af nedenstående universmodel (fig. 1), at oversætteren

1. må besidde en faglig og en sproglig viden på *både dansk og fremmedsproget*, der svarer ikke alene til det specifikke vidensfelt regnskabsprincipper, men også til videnselementer i fagområder, der betragtes som overordnede i forhold til regnskabsprincipper (fx hele regnskabsområdet, herunder internt og eksternt regnskab) samt bl.a. områderne selskabsret (der vedrører virksomhedsformer herunder koncerner), valuta og skat;
2. må besidde stor viden om og indsigt i almensproget på både dansk og fremmedsproget;
3. må kunne anvende den viden, der er anført i 1 og 2.



Figur 1



Figur 2

Oversætteren behøver imidlertid ikke være ekspert i de ovennævnte områder for at løse opgaven tilfredsstillende, det er tilstrækkeligt, at vedkommende besidder et udsnit af denne

viden svarende til de centrale begreber inden for de nævnte områder, hvilket kan illustreres ved dette udsnit ("slipseudsnit") i fig. 2.

Oversætteren skal - for at udtrykke det på en anden måde - kunne gå fra den ene verden (i dette tilfælde den danske) til den fremmedsprogede verden, der ikke nødvendigvis er identisk med den danske. Såfremt oversætteren ikke besidder den netop beskrevne viden, må vedkommende erhverve sig denne viden. Typisk vil oversætteren søge i ordbøger, men spørgsmålet er, om ordbøgerne indeholder informationer, der svarer til brugerens behov. Dette spørgsmål får vi et svar på ved at analysere informanternes færdighed i at forstå den danske fagtekst og oversætte den til hhv. fransk og spansk.

### 3 Foreløbige resultater af analysen

#### 3.1 Analysens omfang

De resultater, vi vil fremlægge, er primært relateret til de problemer, informanterne har registreret i *de to første afsnit* af den omtalte tekst. Analysen er baseret dels på informanternes egne vurderinger vedrørende problemer og løsninger og dels på vores vurdering heraf. Mange informanter angiver usikkerhed på den løsning, de er nået frem til ved at markere "mon det passer", "sikkert helt forkert", "jeg gætter på X", "jeg tager chancen" o. lign. Vi tager dette som et tegn på, at der stadig er problemer dvs. behov for yderligere information, uanset at løsningen rent faktisk er rigtig.

På den anden side har vi i tilfælde, hvor informanterne lader til at være tilfredse med løsningen ved enten ikke at kommentere resultatet eller at markere f.eks. "OK", i mange tilfælde registreret forkerte løsninger, hvilket igen viser os, at der er et behov for yderligere information - uanset at dette altså ikke er erkendt af informanten.

Vi vil i det følgende beskrive, hvilke tendenser vi har fundet undersøgelsen peger på med hensyn til brugerstrategier og problemarter.

#### 3.2 Brugerstrategier

Med hensyn til prioritering af hjælpemidler (se tabel 2) viser vores undersøgelse, at der er en præference for valg af tosproglige ordbøger som første opslagsværk - endda primært almensproglige ordbøger. Denne tendens er umiddelbart overraskende, i betragtning af at det drejer sig om fagsproglig oversættelse.

Forklaringen på præferencen for de almensproglige ordbøger er efter vores mening, at der er for ringe tillid til de eksisterende fagsproglige dansk-franske/spanske ordbøger, og at opslag via tysk og engelsk forlænger søgeprocessen. Desuden indeholder enkelte såkaldt almensproglige ordbøger et vist antal fagord, hvilket især gælder for fransk. Det skal i øvrigt tilføjes, at mange informanter har benyttet langt flere opslag til løsning af konkrete problemer end de tre opslag, vi har medtaget her (jf. nedenstående tabel 3 over antal opslag).

Det ses endvidere, at der ikke er nogen imponerende aktivitet med hensyn til at søge sag-relevante oplysninger i danske leksika, lærebøger m.v., inden fokus rettes mod ækvivalenten.

ORDBØGER			1. opsl.	2. opsl.	3. opsl.	1+2+3	%
Dansk-fransk/spansk	Almenspr.	F	85	25	6	21	
		S	102	25	2	37	
Fransk/spansk-dansk	Almenspr.	F	2	6	7	3	
		S	1	3	2	8	
Dansk-fransk/spansk	Fagspr.	F	46	21	7	13	
		S	43	33	12	25	
Fransk/spansk-dansk	Fagspr.	F	0	0	0	0	
		S	6	11	4	6	
Dansk-dansk	Fagspr.	F	23	11	6	7	
		S	1	1	0	1	
Fransk-fransk/ Spansk-spansk	Almenspr.	F	3	10	4	3	
		S	1	2	2	1	
Fransk-fransk/ Spansk-spansk	Fagspr.	F	22	44	36	19	
		S	3	4	6	4	

ANDRE HJÆLPEMIDLER			1. opsl.	2. opsl.	3. opsl.	1+2+3	%
Lærebøger - Danske		F	10	2	2	3	
		S	-	-	-	0	
Lærebøger - Franske/spanske		F	12	4	4	4	
		S	1	1	1	1	
Autentiske tekster - Danske		F	3	2	3	1	
		S	0	-	-	0	
Autentiske tekster - Franske/sp.		F	4	5	4	2	
		S	8	4	11	7	
Diverse		F	0	1	1	0	
		S	0	1	1	1	

Tabel 2 Anvendte hjælpemidler. Procentangivelserne i den højre søjle er beregnet på grundlag af det totale antal opslag, som er 550 for fransk og 350 for spansk.

Og der er, som det fremgår, ingen væsentlig brug af materiale ud over ordbøger. Vi har dog registreret en tendens til som sidste opslag at konsultere autentiske tekster fra franske og spanske årsberetninger som en slags check af den fundne ækvivalent.

M.h.t. antallet af hjælpemidler, der blev benyttet, viser tabel 3 følgende:

Antal opslag			Anvendte hjælpemidler		
Min./max.	F	1-15	Min./max.	F	6-19
	S	1-10		S	4-18
I gennemsnit pr. problem	F	2,9	I gennemsnit pr. informant	F.	ca. 12
	S	2,1		S	ca. 8

Tabel 3 Antal opslag og hjælpemidler

Tallene for tidsforbruget til løsning af de konkrete problemer og til den samlede oversættelse fremgår af tabel 4.

Tidsforbrug pr. problem			Tidsforbrug totalt		
Min./max. (minutter)	F	2-120	Min./max. (timer)	F	2-18
	S	0,5-135		S	3-21
I gennemsnit pr. problem (minutter)	F	12	I gennemsnit pr. informant (timer)	F	ca. 9
	S	9		S	ca. 9

Tabel 4 Tidsforbrug

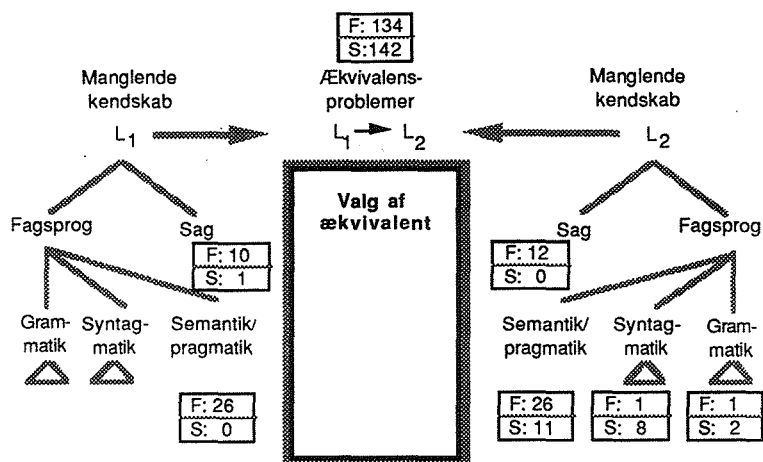
Det skal bemærkes, at tallene for antallet af anvendte hjælpemidler og det samlede tidsforbrug henviser til den samlede oversættelse, i modsætning til alle de øvrige tal, der som nævnt kun refererer til oversættelsen af de to første afsnit.

Som det ses, er der generelt tale om relativt høje tal. Specielt m.h.t. tidsforbruget og antallet af anvendte opslagsværker: 9 timer i snit og 10 opslagsværker i brug for at oversætte en side er efter vores mening ude af proportioner.

### 3.3 Problemarter

Tekstens to første afsnit omfatter ifølge informanternes protokoller i alt 25 problemer. Naturligvis har ikke alle informanter haft samtlige problemer.

De problemer, der umiddelbart kan konstateres ud fra protokollerne, fordeler sig på de problemarter som er vist i figur 3.



Figur 3 Interpretationsmodel

Som det fremgår af interpretationsmodellen, har vi registreret nogle få problemer i forbindelse med reception (= forståelse af teksten), samt nogle få problemer i forbindelse med produktion på L2. Men langt de fleste problemer opstår ved transporten fra L1 til L2 - altså ækvivalensproblemer.

*Ækvivalensproblemet* indebærer imidlertid en række andre problemer af *sproglig* og *faglig* art. Det har ikke været muligt systematisk at dechifrere disse problemer alene på basis af informanternes egne angivelser. Søgning på den rette ækvivalent foregår i ordbøger, encyklopædiske værker, lærebøger og paralleltekster, men uden at det egentligt er muligt objektivt at konstatere, hvad informanten i sine søgestrategier anser for at være det specifikke problem. På basis af sporadiske kommentarer hos informanterne og/eller deres søgeadfærd, altså med fokus på, hvad informanterne siger og hvad de gør, mener vi dog, det er rimeligt at pege på følgende problemtyper, som i tabel 5 er illustreret ved uddrag af protokollerne for den spanske oversættelse.

#### Eks. 1:

Informanten spørger: Hvad hedder *sammendrag*? Vedkommende slår op i en almensproglig dansk-spansk ordbog og finder 4 ækvivalenter. Reaktionen er "ingen specificering, men *resumen* lyder sandsynlig". Vi har kaldt dette og alle analoge problemer:

- Usikkerhed vedrørende ordets semantiske indhold/semantisk differentiering

#### Eks. 2:

I en anden typisk situation spørger informanten i protokollen: Hvad hedder *konsolidering*? Får svaret *consolidación* og spørger igen: Kan jeg bruge det i denne relation? Vi har kaldt dette og lignende problemer:

- Usikkerhed vedrørende ordets placering i universet



**Eks. 3:**

En informant spørger: Hvad hedder *regnskabsposter*? Hun får 3 ækvivalenter ved første opslag og checker dernæst i fagordbogen, men er stadig usikker, går videre til en lærebog og finder den rigtige ækvivalent, Hun vil dog gerne sikre sig, at den nu også bruges i praksis og konsulterer til sidst en autentisk tekst. Vi har kaldt den usikkerhed, der fører til et sådant sidste check:

- Usikkerhed vedrørende ordets anvendelse

**Eks. 4**

Informanterne kan vise usikkerhed med hensyn til, hvad f.eks. konsolidering egentlig er. Vi har kaldt dette og alle analoge problemer:

- Usikkerhed vedrørende "sagen"

**Eks. 5:**

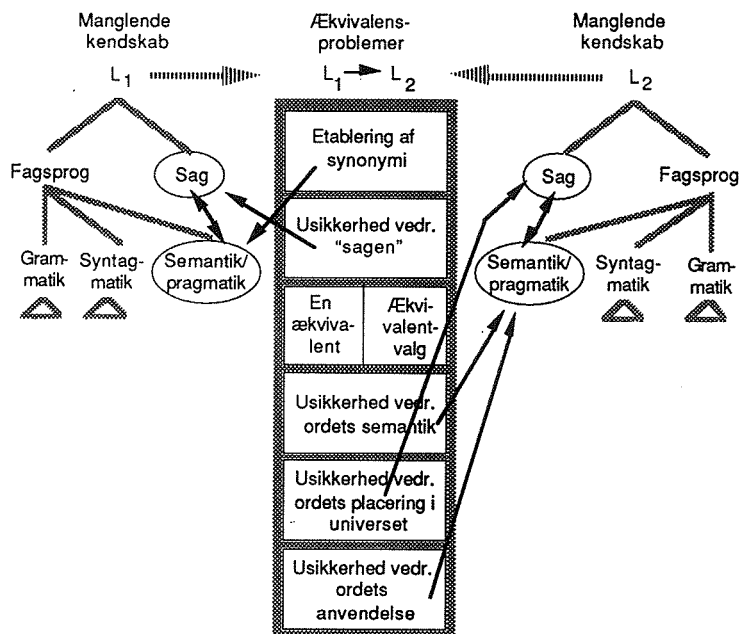
Endelig har vi situationer som den sidstnævnte, hvor man søger en ækvivalent i ordbogen, men ikke finder den, og derfor leder under et andet ord, som man formoder er synonymt med det første. Problemet er her: Er de to danske ord egentlig synonyme?. Vi har kaldt dette:

- Etablering af synonymi

UDDRAG AF PROTOKOL:		
PROBLEM	RESULTAT	INFORMANTENS KOM.
Eks. 1 hvad hedder <i>sammendrag</i> ?	resumen, recolección, compendio, síntesis	"ingen specificering, men <i>resumen</i> lyder sandsynlig"
Eks. 2 hvad hedder <i>konsolidering</i> ?	consolidación	"i denne relation?"
Eks. 3 hvad hedder <i>regnskabsposter</i> ?	1. partida contable, asiento 2. asiento, partida 3. partidas 4. partidas	"checkes" "checkes" + "OK"
Eks. 4 hvad er egentlig <i>konsolidering</i> ?	-	-
Eks. 5 hvad hedder <i>tilbageførelse</i> ?	under tilbageførelse: intet resultat under tilbageføring: retorno	"det må være <i>retorno</i> "

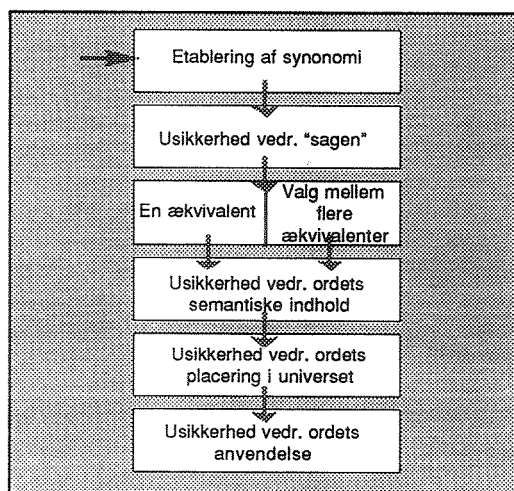
Tabel 5 Uddrag af protokol

Ækvivalensproblemerne har således relation til enten L1 eller L2 i interpretationsmodellen, jf. fig. 4.



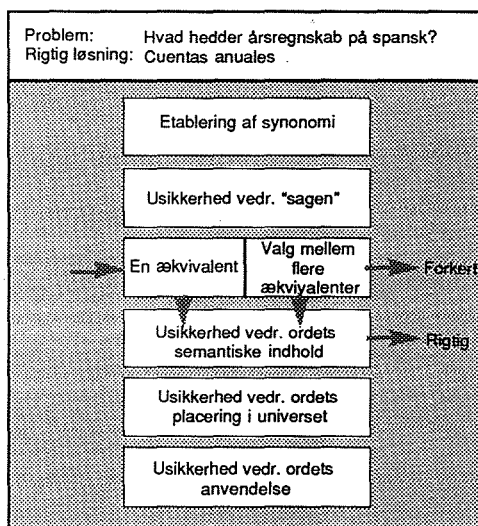
Figur 4 Problemtyper

På basis af de anførte problemtyper har vi i fig. 5 opstillet et flow-diagram ud fra vores opfattelse af, i hvilken rækkefølge det forekommer logisk, at problemerne opstår.



Figur 5 Flow-diagram

Disse problemtyper er i det følgende illustreret med nogle konkrete eksempler på registrerede problemer og vores vurdering heraf. I fig. 6 vises et eksempel på usikkerhed vedr. ordets semantiske indhold.

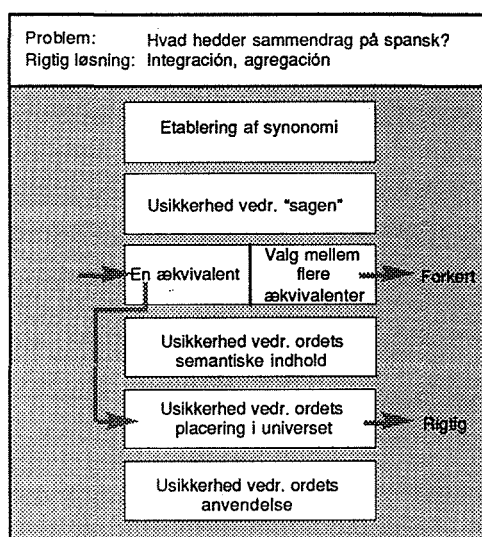


Figur 6 Eks. på usikkerhed vedr. ordets semantiske indhold

Informanterne vælger her ud fra rækken af ækvivalenter i ordbogen typisk *memoria* (årsregnskabs noter), *cuenta de pérdidas y ganancias* (driftsregnskab, resultatopgørelse) eller *balance anual* (balance, status) og er tilsyneladende tilfredse med resultatet. Disse ækvivalenter dækker imidlertid over begreber, der hver især indgår som delelementer af årsregnskabet. Der kan være forskellige forklaringer på deres forekomst som ækvivalent for dette lemma. *Balance anual* har for eksempel tidligere ofte været anvendt som pars pro toto for årsregnskabet som sådan, idet der i Spanien har været tradition for primært at fokusere på balancen ved vurdering af firmaernes økonomiske resultater. Men i denne tekst, hvor tekniske detaljer i alle delelementer af årsregnskabet kommenteres, herunder naturligvis også detaljer i balancen, er termen *balance anual* naturligvis misvisende. Den rigtige ækvivalent er i dette tilfælde *cuentas anuales*. Vi mener, at valget af de forkerte ækvivalenter skyldes usikkerhed omkring de pågældende ækvivalenter semantiske indhold. Hvis brugerne havde erkendt denne usikkerhed og foretaget videre søgninger på det semantiske indhold, ville denne usikkerhed kunne være afhjulpet. Fig. 6 viser den rigtige søgestrategi i den pågældende situation. Det er klart, at leksikografen ville kunne afkorte søgeprocessen ved at forsyne ordbogen med en definition, der ville medføre, at de forkerte løsninger kunne forkastes.

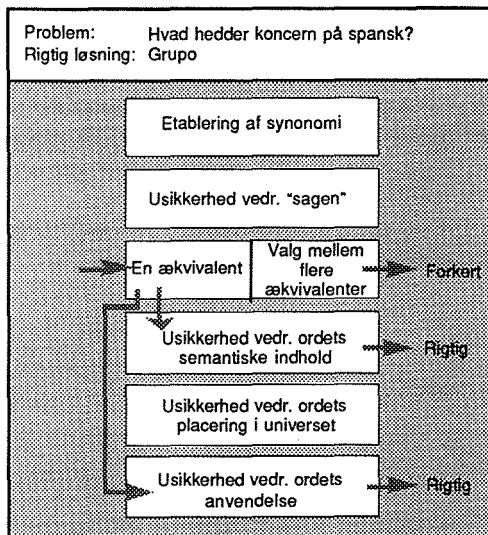
I fig. 7 vises et eksempel, hvor en definition ikke ville være tilstrækkelig til at give en succesfuld bruger. Det drejer sig om oversættelse af termen *sammendrag*. Her vælger informanterne typisk *resumen*. Vi gætter på, at de vælger dette, fordi de går ud fra, at det ækvivalerer med det danske ord *resumé*, (hvilket det rent faktisk også gør), og fordi *sammendrag* er lig med *resumé* i deres almensproglige univers. Men *resumen* - eller *resumé*

- dækker ikke indholdet af *sammendrag* i det univers eller den kontekst, vi befinder os i her. Her betyder det nemlig sammenlægning (af de enkelte årsregnskaber). I det konkrete tilfælde ville en definition af *resumen* således ikke have givet brugeren yderligere information, idet vedkommende sikkert udmærket via den danske ækvivalent kender dette ords betydning. Der er simpelt hen tale om en fejlplacering i universet. En erkendt usikkerhed på dette punkt burde have tilskyndet brugeren til at søge oplysninger om sagen, dvs. i dette tilfælde om konsolideringsprocessen som sådan, hvilket kunne have belyst ordet *sammendrags* homonyme karakter. Fig. 7 viser den rigtige søgestrategi. Leksikografen kunne her hjælpe brugeren med en områdeangivelse.



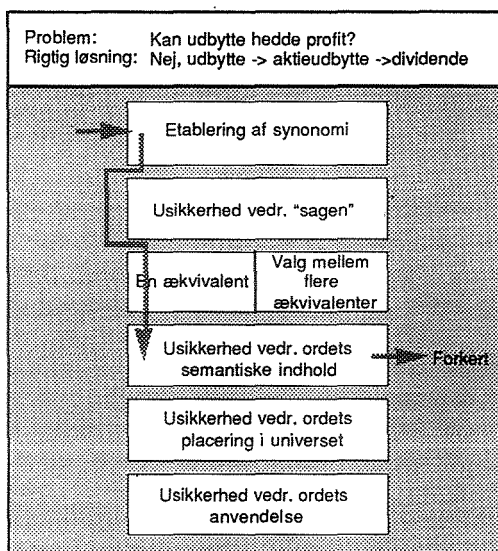
Figur 7 Eks. på usikkerhed vedr. ordets placering i universet

Eksemplet i fig. 8 viser, at brugerne kan have problemer med ordenes anvendelse. Informanterne vælger typisk *consorcio* eller grupo de *empresas*. Valg af *consorcio* tyder på manglende kendskab til ordets semantiske indhold. En afhjælpning af denne usikkerhed ved en god definition ville have givet brugeren den fornødne information om, at *consorcio* dækker over en løsere sammenslutning af virksomheder end en koncern. En definition ville til gengæld ikke have været tilstrækkelig for ækvivalenten *grupo de empresas*. Denne spanske term er egentlig den rigtige ækvivalent, men kun teoretisk, for i praksis bruges kun den korte form *grupo*. Valget af *grupo de empresas* tyder således på en usikkerhed omkring anvendelsen af termen og et autentisk eksempel ville her have givet brugeren den rigtige information.



Figur 8 Eks. på usikkerhed vedr. ordets anvendelse

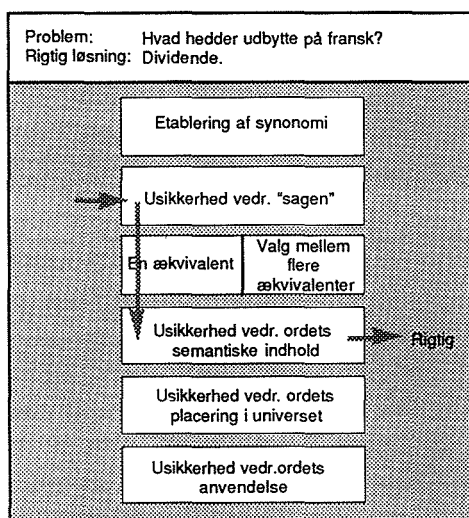
I det følgende vises eksempler på brugerstrategier fra det franske materiale. Eksemplet i fig. 9 illustrerer en strategi, der gav en tilfreds, men dog ikke succesfuld oversætter.



Figur 9 Eks. på fejlagtig etablering af synonimi

Oversætteren undersøger, om *udbytte* og *profit* er synonyme ved hjælp af opslag under *profit* i et dansk-dansk økonomisk leksikon, og etablerer på det grundlag fejlagtigt synonymi (ordbogsartiklen oplyser bl.a., at *profit* = *vinding*, *udbytte*, *fortjeneste*). Brugeren formoder herefter, at det franske *profit* svarer til det danske, men checker for en sikkerheds skyld i en fransk-fransk fagordbog, sammenholder den danske og den franske definition, og oversætter *udbytte* med *profit*, hvilket er forkert. Oversætteren har ikke forstået "sagen", dvs. at der i denne kontekst er tale om *udbytte* = *aktieudbytte*, som hedder *dividende* på fransk.

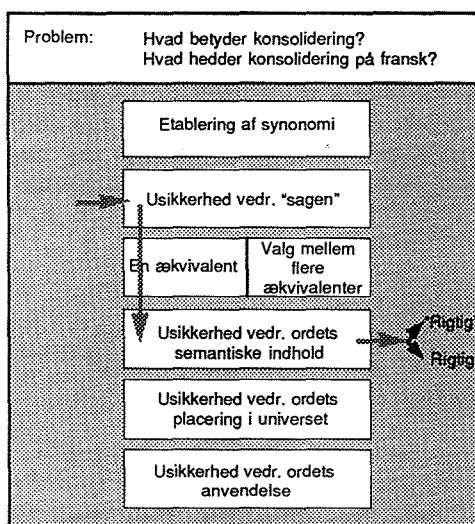
En bedre ordbog kunne have hjulpet denne bruger ved at give informationer om "sagen" og bl.a. oplyse, at *udbytte* har flere betydninger samt evt. henvise til opslagsordene *udbytte* og *aktieudbytte*. Det følgende eksempel i fig. 10 viser en bruger, der p.g.a. en god søgestrategi blev en både tilfreds og succesfuld oversætter.



Figur 10      Eks. på god søgestrategi

I dette tilfælde undersøges "sagen" ved opslag under *udbytte* i en dansk-dansk fagordbog og informanten konkluderer, at *udbytte* = *aktieudbytte*. En fransk ækvivalent findes i en dansk-fransk ordbog uden definitioner, hvorfor ækvivalenten checkes i en fransk-fransk fagordbog. Oversætteren havde i dette tilfælde ikke behøvet det sidste checkopslag, men valgte en sikker og mere tidskrævende strategi. Strategien viser imidlertid, at brugerne - måske som følge af mangel på sikkerhed, om ordbogens forslag passer til den pågældende kontekst - har behov for definitioner på både udgangssprog og målsprog.

I det sidste eksempel følger vi i fig. 11 to brugere, hvoraf den ene *tilfældigt* bliver succesfuld med et korrekt resultat, den anden ligeledes succesfuld men i dette tilfælde p.g.a. en god strategi.



Figur 11 Eks. på rigtig og forkert placering af ord i universet

Begge brugere søger "sagen" oplyst ved at konsultere dansk-danske opslagsværker. Den første (dvs. den uheldige bruger) finder en definition, som vedkommende accepterer. Definitionen dækker imidlertid et begreb inden for et andet vidensfelt, finansiering, hvor konsolidering betyder styrkelse af en virksomheds økonomiske stilling. Den anden bruger finder ligeledes denne definition, forkaster den, søger videre og finder den, der dækker begrebet i denne kontekst, dvs. inden for koncernregnskab, hvor konsolidering betyder en sammenlægning af regnskaberne i en koncerns virksomheder. Begge finder herefter en fransk ækvivalent i en dansk-fransk ordbog, og begge checker det semantiske indhold af den i fransk-franske værker.

Den første bruger finder en definition på det "forkerte" begreb (dvs. inden for finansiering), sammenholder den med den ligeledes "forkerte" danske og slutter - fejlagtigt - at ækvivalenten er korrekt. Bruger nr. 2 finder frem til en definition på det korrekte begreb (dvs. inden for koncernregnskab) og løser dermed problemet korrekt. Den første brugers løsning er imidlertid tilfældigvis også korrekt, da konsolidering på både dansk og fransk hedder det samme, uanset at der er tale om 2 forskellige begreber. Den ene bruger havde således problemer med at få sig placeret i det korrekte univers, hvilket den ideelle ordbog burde hjælpe med ved at give informationer om begge universer.

#### 4 Foreløbige konklusioner

Den foreløbige analyse viser, at det ikke lykkedes brugeren af i gennemsnit ca. 10 hjælpemidler og trods en arbejdsindsats på ca. 10 timer for at oversætte blot 2 afsnit i teksten at komme succesfuldt fra den ene verden over i den anden: generelt blev der ikke produceret en tilfredsstillende oversættelse. Undervejs over til denne anden verden, til dette andet univers, dvs. i oversættelsesprocessen, havde brugeren desuden store, mange og alvorlige problemer.

Brugeren var ofte fortvivlet, usikker og opgivende og råbte om hjælp både i vore protokoller og i interviewene, og dette gjaldt både den grønne og den mere kyndige bruger.

Hvorfor forløb denne oversættelsesproces ikke bedre? Vi mener, det skyldes, at brugerne ikke i tilstrækkeligt omfang beherskede den viden, som vi hævdede var nødvendig for at kunne producere en tilfredsstillende oversættelse. De havde vanskeligt ved at afgrænse tekstens univers, slipseudsnittet, og befandt sig ofte udenfor og dermed i "forkerte" universer (fx finansiering, den dagligdags verden). På grundlag af den første fase af vores analyse mener vi at have afdækket nogle af brugerens problemer og vist, at ækvivalensproblemer udgør langt den største gruppe. Ligeledes har vi vist, at ækvivalensproblemer er en kompliceret størrelse, der bl.a. skyldes

- usikkerhed vedr. sagen
- usikkerhed vedr. ordenes semantiske indhold
- usikkerhed vedr. ordenes placering i universet
- usikkerhed vedr. ordenes anvendelse.

Brugerens behov må derfor være at kunne få løst disse problemer samt at få minimeret antallet af opslag og opslagsværker. Vi kan på basis af ovenstående opstille nogle af de elementer, analysen viser, at den ideelle fagordbog bør indeholde. Opstillingen er i prioriteret rækkefølge i overensstemmelse med ovennævnte problemer:

- information om "sagen" / encyklopædiske oplysninger
- definitioner
- anvendelsesområder (ref. til placering i univers, fx finansiering, regnskab)
- eksempler.

Hvordan disse ønsker bedst kan indbygges i en fagordbog, og hvorvidt det i det hele taget er hensigtsmæssigt at søge alle de nødvendige oplysninger samlet i en og samme bog, har vi ikke på nuværende tidspunkt noget bud på. Det er målet for hele undersøgelsen og vil udgøre den sidste del af projektet. Men vi vil sige ligesom den tyske sprogforsker Hans J. Vermeer (1989: 171):

Je weniger ein Wörterbuch nur Wörterbuch ist, desto hilfreicher ist es für den Translator (altså: jo mindre en ordbog kun er ordbog, desto mere nyttig er den for oversætteren).

I øvrigt vil vi i forbindelse med vores konklusion henvise til artikler i denne publikation af deltagere i det overordnede projekt på Handelshøjskolen i Århus: Sandro Nielsen: *Forholdet mellem alordbøger og enkeltfagsordbøger* og Sven Tarp: *Encyklopædiske oplysninger i tosprogede tekniske ordbøger*. Disse artikler supplerer ovennævnte konklusion.



## Litteratur

- Benbow, Timothy et al. 1990. Report on the New Oxford English Dictionary User Survey. I: *International Journal of Lexicography*. Volume 3, No. 3, s. 155-204.
- Bogaards, Paul. 1990. Où cherche-t-on dans le dictionnaire? I: *International Journal of Lexicography*. Volume 3, No. 2, s. 79-103.
- Diab, Turki. 1990. *Pedagogical Lexicography: A Case Study of Arab Nurses as Dictionary Users*. Tübingen: Niemeyer.
- Neubauer, Fritz. 1987. Auf der Spur des "unbekannten Wesens": Der DaF-Wörterbuchbenutzer. I: *Wörterbücher und ihre Didaktik*. Hrsg. v. E. Zöfgen. Bad Honnef: Keimer/Zürich: Hebsacher (=Themheft der Zeitschrift BBS 14 (1985)), s. 216-235.
- Ripfel, Martha. 1989. Ergebnisse einer Befragung zur Benutzung ein- und zweisprachiger Wörterbücher. I: *Lexicographica* 5, s. 178-201.
- Ripfel, Martha/Wiegand, Herbert Ernst. 1988. Wörterbuchbenutzungsforschung. Ein kritischer Bericht. I: Herbert Ernst Wiegand (ed.): *Studien zur neuhochdeutschen Lexikographie VI*, 2. Teilband. Hildesheim: Olms Verlag, 491-520.
- Vermeer, Hans J. 1989. Wörterbücher als Hilfsmittel für unterschiedliche Typen der Translation. I: Franz Josef Hausman/Oskar Reichmann/Herbert Ernst Wiegand/Ladislav Zgusta (eds.): *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography*, First Volume. Berlin, New York: de Gruyter, 171-174.
- Wiegand, Herbert Ernst. 1985. Fragen zur Grammatik in Wörterbuchbenutzungsprotokollen. Ein Beitrag zur empirischen Erforschung der Benutzer einsprachiger Wörterbücher. I: Henning Bergenholtz/Joachim Mugdan: *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch* 28.-30.6.1984. Tübingen: Niemeyer, 20-98.



Del 3



Sven-Göran Malmgren

## Om behandlingen av polysema ord i enspråkiga ordböcker

I uppsatsen tas två aspekter på behandlingen av polysema ord i enspråkiga ordböcker upp: indelningen i betydelsemoment (*lexem*) och (mer kortfattat) metaspråkliga markörer för att ange relationen mellan huvudbetydelse och "avledda" betydelser (inom ett *lexem*). Traditionella kriterier för lexemindelning (formella, syntagmatiska och paradigmatiska enligt Svensén, *Handbok i lexikografi*) diskuteras jämte några specialkriterier från den nyare litteraturen. Dessutom lyfts begreppet *reguljär polysemi* fram, och det föreslås att ord som tillhör samma polysemityp (*analogt polysema ord*) behandlas på samma sätt, vad lexemindelningen beträffar.

Ett av de viktigaste och mest intrikata problemen inom enspråkig lexikografi är hur artiklar om mångtydiga ord ska struktureras. Även om vi antar att man på något sätt löser problemet med gränsdragningen mellan homonymi och polysemi, en fråga som jag inte tänker ta upp här, återstår - det enligt min mening betydligt större - problemet: hur behandla polysemi? Vilken betydelse ska anses vara huvudbetydelse, i vilken ordning ska överförda, utvidgade etc. betydelser ges, när ska man upprätta flera numrerade betydelsemoment och när bara ett, vilka metaspråkliga markörer kan man använda för att antyda sambandet mellan de "avledda" betydelseerna (termen från Andersson 1977, s. 95) och huvud- eller kärnbetydelsen - se där några av de problem lexikografen ständigt ställs inför. I den här uppsatsen kommer de båda sist nämnda problemen, framför allt indelningen i betydelsemoment, att diskuteras. Utgångspunkten är ett avsnitt i Bo Svenséns *Handbok i lexikografi*, men även andra bidrag till den teoretiska diskussionen kommer att beaktas. En typ av kriterium, som såvitt jag vet inte övervägts tidigare i den svenska debatten, kommer att föreslås. Också praktiken i enspråkiga ordböcker, främst *Svensk ordbok* (som jag har medarbetat i), kommer att beröras något.

### Om indelning i betydelsemoment

Är det viktigt för en lexikograf att ta ställning till frågan om/när en viss betydelsenyans, avvikande från en på något sätt etablerad huvudbetydelse, ska behandlas som ett eget (numrerat) betydelsemoment? Man kan möta uppfattningen att detta mest är en praktisk fråga (se t.ex. Eriksson 1975, med kritik av denna uppfattning). Enligt min mening är det emellertid viktigt att lexikografen gör ett medvetet val, baserat på en gång fastlagda principer för ordboken ifråga, även om man naturligtvis inte alltid kan hävda att en viss lösning är den bästa tänkbara. Låt oss utgå från ett enkelt exempel, verbet *bannlysa*. Man kan tänka sig två utformningar av en (fiktiv) ordboksartikel över detta ord, byggda på exakt samma lexikografiska material:

A. *bannlysa* ... utesluta ur (katolska) kyrkans gemenskap - äv. utvidgat: utesluta ur gemenskap - äv. ytterligare utvidgat: förbjuda förekomst av

B. *bannlysa* ... 1. utesluta ur (katolska) kyrkans gemenskap  
2. utesluta ur gemenskap 3. förbjuda förekomst av

(De båda senare betydelseerna kan illustreras med språkproven *den oppositionella gruppen bannlystes på partikongressen* resp. *Runda ord var tidigare bannlysta i Svenska Akademiens ordlista*.)

Jag skulle vilja hävda att dessa båda tänkta ordboksartiklar åtminstone teoretiskt bör implicera två olika påståenden om (språkbrukarnas kunskap om) ordet *bannlysa*. Den första beskrivningen bör innebära att det - enligt den tänkte artikelförfattaren - också synkront finns en nära relation, även i språkbrukarnas "abstrakta" lexikon, mellan de båda avledda betydelseerna och huvudbetydelsen. Denna relation manifesteras bl. a. genom att en språkbrukare som hör satsen *runda ord var tidigare bannlysta...* associerar till huvudbetydelsen av *bannlysa*. Enligt den andra artikelversionen finns ingen relation av detta slag.

En ordbok som huvudsakligen ordnar artiklarna över polysema ord enligt den första principen kan sägas ha en *hierarkisk struktur*, medan en ordbok som snarare bygger på den andra principen har en *linjär struktur* (Werner 1989). Gränsen mellan hierarkisk och linjär struktur är dock flytande; dels kan relationer mellan numrerade betydelsemoment anges, dels kan dessa relationer vara implicita och t. ex. framgå av definitionerna. Ett grovt mått på hierarkiseringen i en ordbok är förhållandet mellan antalet numrerade betydelsemoment och antalet artiklar. Ju lägre (dvs. ju närmare 1) detta relationstal är, desto mer hierarkisk bör ordbokens struktur vara. För *Svensk ordboks* del är detta tal strax under 1,15 (ca 66000/58000; se Malmgren 1988), vilket troligen är lågt. Ett par exempel på ordböcker med betydligt mer linjär struktur torde vara *Longman Dictionary of Contemporary English* och, för svenskans del, *Illustrerad svensk ordbok*.

I arbetet på *Svensk ordbok* användes termen *lexem* praktiskt i betydelsen 'numrerat betydelsemoment'. Teoretiskt borde ett lexem vara en väl avgränsad (del)betydelse av ett (polysemt) ord, jämte närliggande betydelser, avledda enligt vissa, i möjligaste mån specificerade, regler. (Jfr Järborg 1984, s. 97 ff.) Termen *lexem* (eller motsvarande term på andra språk) används emellertid inte ofta i denna betydelse; betydligt vanligare är *semem* eller *betydelsemoment* (eller motsvarande). I denna uppsats kommer jag dock huvudsakligen att använda terminologin från arbetet på *Svensk ordbok*.

I sin *Handbok i lexikografi* anger Bo Svensén fyra typer av kriterier för avgränsning och identifiering av delbetydelser av polysema ord: formella, syntagmatiska, paradigmatiska och pragmatiska (Svensén 1987, s. 106-109). Svensén ger bl.a. följande exempel:

Formella kriterier:	<i>finans</i>	1. (vanl. plur.) 2. (vanl. sing.)
Syntagmatiska kriterier:	<i>flytta</i>	1. (transitivt) 2. (intransitivt)
Paradigmatiska kriterier:	<i>säkerhet</i>	1. (syn. <i>trygghet</i> ) 2. (syn. <i>visshet</i> )
Pragmatiska kriterier:	<i>kasta</i>	3. (sömnadsterm) 4. (veterinärspråk)

Den första typen av kriterier har fördelen att vara handfast och enkel att tillämpa. Troligen avser Svensén huvudsakligen ordböjning, men i vissa fall borde också ordbildningen kunna komma in i bilden. Ett argument för att räkna med två lexem *stor* ('omfångsrik' resp. 'betydande') kunde sålunda vara att den förra betydelsen motsvaras av substantivet *storlek*, den andra av substantivet *storhet*.

Den andra typen av kriterier är likaså viktig, men kanske tillämpas den ibland väl mekaniskt. I sin kritik av den berömda tredje upplagan av *Websters* ordbok skriver Weinreich (1964), på tal om just specialfallet transitiv/intransitiv användning, att ordboken överanvänder detta kriterium, vilket enligt Weinreich "is not only out of keeping with the productiveness of detransitivizing patterns in English, but also leads to an uneconomical repetition of definitions". Weinreichs argument och kritik torde utan vidare vara tillämpliga också på svenskan och på vissa svenska ordböcker.

Den tredje typen av kriterier är jämte den andra den som främst framhävs av Bengt Loman i hans genomgång av metoder för särskiljning av betydelser i arbetet på *Svenska Akademiens ordbok*, SAOB (det s.k. utbytesprovet; Loman 1965, s. 49 ff.). Loman ger exempel på såväl "positiv" som "negativ" användning av kriteriet. Som argument för att upprätta tre lexem *retsam* i SAOB, motsvarande de betydelsenyanser som exemplifieras av fraserna *han är mycket retsam av sig*, *varför skulle han ta sig en sådan retsam överlägsen ton mot henne?* och *oturen kom att spela honom ett retsamt spratt*, anger Loman att *retsam* i första fallet kan bytas ut mot *kivaktig*, i det andra mot bl. a. *förarglig* och i det tredje mot bl. a. *irriterande*. Å andra sidan bör man enligt Loman i nusvenskan inte upprätta två lexem *god* som motsvarar användningen i fraserna *var god och hjälp mig* resp. *en god människa*. Ännu i 1800-talssvenskan kunde man i det första - men inte i det senare - fallet byta ut *god* mot synonymen *artig*, men i nusvenskan är detta inte möjligt - detta är motiveringen. Enligt min mening är det dock tveksamt om man ska använda utbytesprovet "negativt"; det är ju, som ibland har påpekats, delvis en ren tillfällighet om en viss synonym existerar eller inte. I fallet *god* kunde man väl hävda, att det finns en skillnad mellan *god* i *var god och hjälp mig* och *god* i *en god människa* - den nämligen, att *god* i det förra fallet betecknar en tillfällig, i det senare fallet en varaktig egenskap. Detta är visserligen vanligt bland adjektiven, men just hos adjektivet *god* torde den tillfälliga innebörden vara sällsynt. (Jfr även nedan, beträffande *retsam*.)

Över huvud taget är det en viktig fråga om Svenséns kriterier kan användas både positivt och negativt. Det är naturligtvis ingen tvekan om att ett positivt utslag av ett kriterium innebär ett visst - men inte nödvändigtvis tillräckligt - argument för lexemuppdelning, men det är alls inte lika självklart att ett negativt utslag innebär ett argument mot. Ett problem är också, att kriterierna, om de får användas både positivt och negativt, kan komma i konflikt med varandra - vilket ska då väga tyngst?

En genomgång av den internationella litteraturen på området (i första hand de artiklar som förtecknas under rubriken "Polysemy" i Zgusta 1988) ger inte så mycket utöver Svenséns kriterier. Några intressanta förslag kan dock noteras. Ayto (1983) diskuterar betydelseindelningen av det engelska ordet *cup* på grundval av en semantisk fältanalys. Det gäller dels den betydelse som i stort sett motsvarar svenskans *kopp*, dels en betydelse som vanligen skulle återges med *(plast)mugg* på svenska (ett ej kupigt dryckeskärl av plast, utan fat, vanligen använt till varma drycker). Låt oss kalla dem *cup/1* och *cup/2* utan att "1" och "2" därför behöver innebära "lexemstatus". Bör man upprätta två lexem för engelskans del? Ayto ger inget definitivt svar, men ett av hans argument för en eventuell tvålexemlösning är intressant. I det semantiska fält där *cup/1* och *cup/2* ingår är den närmaste "grannen" till *cup/1 mug* och

den närmaste grannen till *cup/2 glass*. Om man anser, att avgränsningen av *cup/1* gentemot *mug* och avgränsningen av *cup/2* gentemot *glass* behöver framhävas tydligt, då görs detta enligt Aytos bäst genom att de båda betydelseerna behandlas under skilda numererade betydelsemoment. Man kan säga att Aytos argument innebär en sorts generalisering av SAOB:s utbytesprov (jfr ovan).

I en stor uppsats har Igor Melčuk (1988) nyligen diskuterat de flesta aspekter på den semantiska beskrivningen i den typ av ordbok som har Melčuk m. fl. som upphovsmän, bl.a. kriterier för lexemindelning (Melčuk använder f. ö. termen *lexem* på i stort sett samma sätt som den används i denna uppsats). Ett av dessa kriterier är av speciellt intresse även om det liksom Aytos kan sägas falla in under Svenséns tredje grupp. Det enklaste specialfallet kan här illustreras med hjälp av ett något övertydligt svenskt exempel, verbet *laga* i de båda betydelser som illustreras av fraserna *laga mat* resp. *laga bilen*. Om man samordnar de båda fraserna kan man få en sats av typen *hon lagade både maten och bilen*, en sats som verkar något underlig och förmodligen skulle uppfattas som ett skämt eller som en stilistisk figur, ett zeugma. Detta tyder på att vi bör räkna med två lexem i detta fall. Allmänna: ju mer avvikande eller humoristiskt resultatet av liknande samordningstest blir, desto större skäl har man att räkna med två lexem.

Intressant och originellt är också ett förslag av Ilse Karl (Karl 1983). Karl diskuterar behandlingen av verbet *laden* och ställer sig frågan om de båda betydelsenyanser som förekommer i satserna *Die Schauerleute laden die schweren Kisten auf das Schiff* resp. *Er lädt den Kranken auf seine Schultern* ska ge upphov till två skilda lexem. Hennes svar är ja, med motiveringen att det är den sista, och enbart den sista, betydelsenyansen som utgör utgångspunkt för överförda betydelser av typen *Ich lade eine grosse Verantwortung auf meine Schultern*. Dvs. förekomst av överförda betydelser och deras relationer till två givna "egentliga" betydelser bör enligt Karl påverka lexikografens beslut huruvida dessa egentliga betydelser ska betraktas som tillhörande ett eller två lexem.

Allmänt kan det sägas, att en inte ovanlig kritik av ordböckers artiklar om polysema ord är att de är för lite hierarkiska, dvs. de upprättar för många lexem. Utom Weinreich (1964) kan i det sammanhanget nämnas Lenoble-Pinson (1985) och Moon (1987). Detta är, som framgår av det föregående, helt i linje med tankarna bakom *Svensk ordbok*.

Vi övergår nu till en idé som i motsats de föregående knappast går att passa in i Svenséns schema. Den bygger på ett arbete av den sovjetiske lingvisten Jurij Apresjan (Apresjan 1980) och har i princip, med någon liten modifiering, presenterats tidigare i Malmgren (1988). Här ska jag utveckla den lite vidare och ge den en mer lexikografisk vinkling.

Med *reguljär polysemi* menar jag i anslutning till Apresjan (1980, s. 243 ff.) det förhållandet att en och samma polysemityp förekommer vid minst två icke-synonyma ord i ett språk. Man kan också tala om *starkt* och *svagt produktiv (reguljär) polysemi*, beroende på om polysemitypen förekommer vid många eller få av de ord där den är teoretiskt möjlig. (Här avviker mitt språkbruk något från Apresjans.) De ord som ingår i en ordgrupp som karakteriseras av regulär polysemi kan kallas *analogt polysema*.

Exempelvis är den typ av polysemi (1) som förekommer vid orden *talang*, *kapacitet* och *skönhet* (i samtliga fall 'egenskap (E)' och 'person med egenskapen E') uppenbarligen reguljär men förmodligen ganska svagt produktiv (ord som *klokhets*, *sensibilitet* m.fl. betecknar enbart egenskaper). En annan polysemityp (2), exemplifierad av fraserna *en klok person* - *ett klokt svar* ('som har egenskapen E' - 'som vittnar om egenskapen E'), är mycket starkt



produktiv. Svagt produktiv är däremot typen (3) 'som har egenskapen E' - 'som orsakar egenskapen E' (*en förtvivlad person - en förtvivlad situation*).

Den närliggande tanke som jag skulle vilja hävda är nu att analogt polysema ord bör behandlas på samma sätt i en ordbok, vad lexemindelningen beträffar, och att man har mindre skäl att upprätta två lexem, ju starkare produktiv polysemitypen i fråga är. Dessa kriterier kan kallas *analogikriteriet* resp. *produktivitetskriteriet*. Naturligtvis är analogikriteriet endast relativt; det förutsätter ett avgörande grundat på andra kriterier, t. ex. produktivitetskriteriet. Enligt detta senare kriterium bör typen (2) inte leda till lexemuppdelning, medan typen (3) borde göra det. Typen (1) är mer tveksam, men hur som helst gäller analogikriteriet. Lomans exempel *retsam* ovan (de båda första betydelserna) är av typ (2). Här kommer sålunda utbytesprovet (som ger argument *för* lexemuppdelning) i konflikt med produktivitets- och analogikriteriet (som ger argument *mot* lexemuppdelning). Själv skulle jag vilja hävda att i sådana fall de båda senare kriterierna normalt bör ges företräde. Man kan slutligen observera, att produktivitetskriteriet ligger väl i linje med citatet av Weinreich ovan.

### Om angivande av relationerna mellan betydelserna hos polysema ord

Bör man ange relationerna mellan huvudbetydelsen av ett lexem och de avledda betydelserna i en ordbok? Enligt min mening är svaret ja; i synnerhet bör detta gälla ordböcker med hierarkisk struktur, bl. a. som en sorts motivering till att betydelseskilnaderna inte gett upphov till flera lexem. Ett problem är emellertid att det lexikografiska metaspråket är föga utvecklat, och dessutom måste de lexikografiska termer som används i en ordbok ha en relativt allmänspråklig karaktär. Frågan är t. ex. om en egentligen oundgänglig term som *metonym(isk)* kan användas i en svensk ordbok.

I *Svensk ordbok* används termer som *överfört*, *bildligt*, *utvidgat*, *speciellt*, *med konstruktionsväxling* (de tre senare specialfall av *metonymi*, en term som inte användes). Vid många typer av avledda betydelser kan dock relationerna till resp. huvudbetydelse inte anges med någon av dessa termer. Ett exempel bland många är typen 'verksamhet (V)' - 'personer som ägnar sig åt V' (t. ex. *ledning*). I sådana fall lämnas normalt ingen relationsuppgift i *Svensk ordbok*. Allmänt sett torde man dock kunna säga att det lexikografiska metaspråket är jämförelsevis rikt i denna ordbok. Undersöker man enspråkig lexikografi över en del andra språk än svenska får man ett starkt intryck av att särskilt de franska ordböckerna flitigt ger metaspråkliga markörer. Ett gott exempel är *Trésor de la langue française*. Vi kan se på ett utdrag ur artikeln *panache* i denna ordbok (exemplet efter Gorcy 1989):

- A. Grande plume ou bouquet des plumes /.../  
 1. Dans les coiffures des femmes/.../  
 2. Dans les coiffures masculines à certaines époques /.../  
 - *En partic.* Signe distinctif porté par un chef militaire  
 /.../  
 P. *méton.* L'uniforme militaire (de parade) /.../

- B. *P. anal.* (avec la forme, la légèreté, l'éclat d'un panache)  
 /.../  
 - *Spécialement*  
 ARCHIT.  
 HORTIC.  
 /.../
- C. *Au fig.* (P. réf. à la mode d'orner de plumes les casques  
 des guerriers /.../)

Utöver den mycket långt drivna hierarkiseringen lägger man särskilt märke till relationsangivelsen "par méton." och den ganska subtila skillnaden mellan "P. anal." och "Au fig.". Troligen förutsätter ett så avancerat metaspråk emellertid en ganska hög ordbokskultur, en vana hos användarna att umgås obehindrat med ordböcker. Dessa förutsättningar är otvivelaktigt uppfyllda i Frankrike (se särskilt Hausmann 1983) men kanske inte i lika hög grad i Sverige och Norden. Trots att "Trésor-modellen" förefaller utomordentligt tilltalande kan den därför naturligtvis inte utan vidare övertas vid beskrivningen av andra språk. Men nog borde den kunna ge impulser bl. a. åt nordisk lexikografi!

### Litteratur

- Andersson, Erik. 1977. *Verbfrasens struktur i svenskan*. Åbo.
- Apresjan, Jurij D. 1980. *Semantyka leksykalna*. Warszawa. (Polsk övers. av det ryska originalet.)
- Ayto, John. 1983. On specifying meaning. I: Reinhard R.K. Hartmann, red. *Lexicography: principles and practice*: 89-98. London.
- Eriksson, Ulrik. 1975. Om betydelse, betydelseanalys och betydelsebeskrivning. I: *Kring en ordbok. Festskrift till Sven Ekbo 7 augusti 1975*: 1-45. Stockholm.
- Gorcy, Gérard. 1989. Différenciation des significations dans le dictionnaire monolingue: problèmes et méthodes. I: Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta, red. *Wörterbücher/(Dictionaries/Dictionnaires). Ein internationales Handbuch zur Lexikographie (/.../). Erstes Teilband (/.../)*: 905-917. Berlin & New York.
- Hausmann, Franz Josef. 1983. Wörterbücher in Frankreich und Deutschland. Ein Vergleich. I: Herbert Ernst Wiegand, red. *Studien zur neuhochdeutschen Lexikographie III* (Germanistische Linguistik 1-4/82): 119-156. Hildesheim, Zürich & New York.
- Illustrerad svensk ordbok*. 1963. 3. uppl. Stockholm.
- Järborg, Jerker. 1984. *A semantic model for a performance framework. Studies in the possibilities of formalization*. Göteborg.
- Karl, Ilse. 1983. Zur Abgrenzung von Sememen. I: Joachim Schildt & Dieter Viehweger, red. *Die Lexikographie von heute und das Wörterbuch von morgen: Analysen - Probleme - Vorschläge*: 25-36. Berlin.
- Lenoble-Pinson, Michèle. 1985. Polysémie et lexicographie: Quelques dictionnaires descriptifs du français contemporain ... et un dictionnaire encyclopédique. I: René Jongen, red. *La polysémie: Lexicographie et cognition*: 81-93. Louvain.
- Loman, Bengt. 1965. Betydelseanalys. I: Sven Ekbo & Bengt Loman. *Vägledning till Svenska Akademiens ordbok*: 36-61. Stockholm.

- Longman Dictionary of Contemporary English*. 1978. London.
- Malmgren, Sven-Göran. 1988. On regular polysemy in Swedish. I: *Studies in computer-aided lexicology*: 179-200. Stockholm.
- Melčuk, Igor. 1988. Semantic Description of Lexical Units in an Explanatory Combinatorial Dictionary: Basic Principles and Heuristic Criteria. I: *International Journal of Lexicography* 1:3: 165-188.
- Moon, Rosamund. 1987. Monosemous Words and the Dictionary. I: Anthony P. Cowie, red. *The Dictionary and the Language Learner: Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985* (Lexicographica, series maior, vol. 17): 173-182. Tübingen.
- SAOB = *Ordbok över svenska språket utgiven av Svenska Akademien*. 1898-. Lund.
- Svensén, Bo. 1987. *Handbok i lexikografi*. Stockholm.
- Svensk ordbok*. Utarbetad vid Språkdata, Göteborgs universitet. 1986.
- Trésor de la langue française. Dictionnaire de la langue du 19ème et du 20ème siècle*. 1971-85. Paris.
- Weinreich, Uriel. 1964. Webster's third: a Critique of its Semantics. I: *International Journal of American Linguistics* 30 (1964): 405-409.
- Werner, Reinhold. 1989. Probleme der Anordnung der Definitionen im allgemeinen einsprachigen Wörterbuch. I: Franz Josef Hausmann etc., red.: 917-930. (Se Gorcy 1989.)
- Zgusta, Ladislav (with the assistance of Donna M.T.Cr. Farina). 1988. *Lexicography Today. An annotated bibliography of the theory of lexicography*. (Lexicographica, series Maior, vol. 18.) Tübingen.

## Om definitioner i almensproglige ordbøger

Der er i dag en stigende interesse for definitionen i almensproglige et-sprogs-ordbøger. Udviklingen inden for edb har dels ført til nye krav til ordbogsmaterialet, dels til nye muligheder for dets tilrettelæggelse. I den forbindelse har man set på, om resultater inden for sprogfilosofi, lingvistik og cognitive science kan bidrage til en forbedring og nyttiggørelse af ordbogsdefinitionen. Med baggrund i Hilary Putnam's teori om stereotypier vil der i artiklen blevet foreslået to regler for god definitionspraksis i ordbøger.

### 1. Indledning

Nedenstående artikel vil behandle ordbogsdefinitionen og dens form set i lyset af den fornyede internationale debat om definitionens rolle i almensproglige et-sprogs-ordbøger.

Efter først at have set på, hvilke muligheder, der i fremtiden vil være for definitionsarbejdet i ordbøger (afsnit 2), vil jeg efter en diskussion af det traditionelle definitionsbegreb (afsnit 3) med baggrund i den amerikanske filosof Hilary Putnam's teori om stereotypier (afsnit 4) forsøge at udlede nogle retningslinier for god definitionspraksis (afsnit 5).

Jeg vil give retningslinierne form af to regler, som jeg derefter vil kommentere og belyse via nogle eksempler.

Reglerne er skitseagtigt formuleret og gør på ingen måde krav på at være udtømmende. De skal snarere ses som en demonstration af, at det er muligt at anvende overvejelser og resultater inden for aktuel sprogfilosofi og lingvistik i praktisk ordbogsarbejde.

### 2. Nye muligheder for systematisk betydningsbeskrivelse i ordbøger

Der har i de senere år været stigende interesse for definitioner i ordbøger. Det skyldes ikke mindst de muligheder, edb giver for at anvende definitioner inden for edb-områder som automatisk indeksering, naturligsprogsinterfaces, ekspertsystemer, grammatikrettere, automatisk oversættelse, m.m.

Indtil videre har man, når man ser bort fra nogle få undtagelser som fx Cobuild, overvejende koncentreret sig om at anvende edb på definitioner i eksisterende ordbøger. Det gælder fx Calzolari (Calzolari 1990), Boguraev og Briscoe (Boguraev & Briscoe 1990) og Byrd (Byrd et al. 1986). Walker (1989) er en kort, nyere oversigt over tidligere, igangværende og påtænkte projekter.

Resultaterne er ikke uden interesse. Blandt andet er en thesaurus udviklet af Byrd på baggrund af et maskinelt udtræk fra definitionerne i *Webster (7th edition)*. På den anden side kan resultaterne selvfølgelig ikke blive bedre, end der er baggrund for i selve de anvendte ordbøger, der jo oprindeligt har været anlagt ud fra helt andre formål.

Derfor er det, ud over brugen af eksisterende ordbogsmateriale også så småt sat forskning i gang med henblik på en nyudvikling af definitionen og en generel opstramning

af ordbogsdefinitionens indhold og form. Et eksempel herpå er CompLex (IBM research center, Yorktown USA), ledet af Byrd (cf. Byrd 1989).

Men ønsket om en opstramning af definitionens indhold og form og i det hele taget den semantiske struktur i ordbøger eksisterer ikke blot inden for den informationsteknologiske gren af ordbogsforskningen. Ønsket eksisterer efterhånden bredt blandt leksikografer. Således er der i de seneste år opstået et helt nyt forskningsfelt blandt sprogfilosofier, lingvister, kognitive psykologer og leksikografer, blandt andet med baggrund i de muligheder, den edb-mæssige tilrettelægning af ordbogsarbejdet giver. Med mulighed for opbygning af *multifunktionelle baser*, er der groet en erkendelse frem af, at der er behov for en teoretisk nyjustering og underbygning af den hidtidige forskningspraksis, så flere anvendelser end de traditionelle kan tilgodeses.

Hudson (1988) kan ses som et eksempel på opfordring til integration af moderne sprogfilosofi og lingvistik i leksikologi. Som et af de første teoretiske hovedværker inden for dette nye forskningsfelt kan nævnes Eikmeier & Rieser (1981).

Andre artikler og værker er udsprunget af mere praktiske erfaringer med edb i ordbogsarbejde. Sinclair et al. (1986), der er et slags manifest over arbejdet med Cobuild (cf. ovenfor), afsætter hele fire kapitler af deres bog specielt til diskussionen af definitionens opbygning.

Et andet værk, ligeledes udsprunget af et praktisk ordbogsprojekt, men alligevel af en mere teoretisk karakter end Sinclair, er Harras et al. (1991). Bogen, der udelukkende omhandler ordbetydningen og dens fremstilling i ordbøger, kan - for de to første artiklers vedkommende - ses som en redegørelse for de seneste års udvikling inden for lingvistikken i synet på ordbetydning og leksikon.

Det, jeg har at sige i det følgende, kan ses i forlængelse af de nævnte bestræbelser på at forbedre definitionen i ordbøger. Jeg vil indlede med det klassiske begreb om definitionen og derefter udsætte det for en kritisk prøvelse via Putnam's stereotypiteori. Man vil se, at selv om den klassiske definition ikke vil kunne indfri alle de løfter, man traditionelt har tillagt den, så lader den sig stadig anvende, omend ud fra en ændret forståelse af den.

### 3. Det klassiske definitionsbegreb

Definitionsbegrebet bygger ifølge traditionel opfattelse på to vigtige teoretiske antagelser om sprog, nemlig antagelsen eller princippet om *synonymi* og antagelsen eller princippet om *meningskompositionalitet*.

Ifølge den første antagelse består der i sproget omfattende sæt af *ensbetydende størrelser*. Er der tale om ord, kalder man dem normalt for *synonymer* (til hinanden). Synonymer er altså defineret som ensbetydende ord i sproget. Således er ifølge gængs antagelse "angst" *synonym til* "frygt" og "frygt" *synonym til* "angst".

Definitioner af synonymi kan gives en *intensionel* udformning, som den netop angivne, hvor der uden omsvøb og direkte tales om betydnings- eller begrebsindhold (jf. "ensbetydende størrelser"). Eller den kan gives en dertil svarende *ekstensionel* udformning, hvor synonymien specificeres i termer af det begrebsomfang, de to synonyme udtryk har, dvs. i termer af de genstande og forhold, de refererer til i verden. Fx kan synonymien mellem "snaps" og "dram" defineres enten ved, at man specificerer det fælles meningsindhold (eksempelvis gennem ækvivalente beskrivelser: en snaps er en drik, *der X, Y, Z ...*, jf. senere) - det er den

intensionelle definition - eller ved, at man sikrer sig, at de refererer til samme fyldte glas i verden - det er den ekstensionelle definition.

I mere helstøbt form lyder den intensionelle definition af synonymi: to udtryk er synonyme, hvis de har samme mening. Og tilsvarende for den ekstensionelle definition: to udtryk er synonyme, hvis de kan erstatte hinanden i alle sprogets sætninger, uden at forandre disse sætningers sandhedsværdi, dvs. ændre sætningerne, så de bliver falske, hvor de før var sande, eller omvendt. Det kan også formuleres kortere, men måske også mere upræcist som: to udtryk er synonyme, når de i alle kontekster refererer til de samme genstande og forhold i verden.<sup>1</sup>

Synonymer kan danne grundlag for *simple synonymidefinitioner*, fx følgende simple ordbogsdefinition fra Nudansk Ordbog<sup>2</sup>:

**liberal ... frisindet, fordomsfri ...**

Normalt findes der også andre former for definitioner i en ordbog end den simple synonymidefinition. For at redegøre for dem, må vi inddrage den anden vigtige antagelse, der indledningsvis blev nævnt, nemlig princippet om meningskompositionalitet (der kan føres tilbage til den tyske logiker Frege).

Ifølge denne består sproget ikke blot af enkeltord, men også af mere omfattende helheder, hvor flere enkeltord er i stand til at danne en ny betydning (i kraft af visse dannelses- eller sammensætningsregler). For at betydningshelheden skal være kompositionelt sammensat må man kræve, at den udelukkende er dannet ud fra de betydningselementer, den består af, plus reglerne for deres sammensætning.

Blandt visse af disse helheder kan vi nu igen finde synonymi. Således er "ræk mig saltet" synonymt med "vil du række mig saltet" og "radikalisme" synonymt med "radikal anskuelse".

Er den ene af de to størrelser, der sammenholdes, et enkeltord, som i det sidste tilfælde, har vi med, hvad vi kunne kalde *standard-ordbogsdefinitionen* at gøre. Det sidste eksempel er da også hentet fra en ordbog, nemlig *Svensk ordbok* (men oversat til dansk). Et andet eksempel (igen fra *Nudansk Ordbog*) er:

**marinere ... nedlægge sild o.l. i marinade...**

Den simple synonymidefinition ved synonymer og definitionen ved synonymi plus kompositionalitet er altså to eksempler på definitioner i ordbøger.

Hvad den sidste type definition angår kan den igen underinddeles i forskellige typer, hvoraf langt den bedst kendte - og mest benyttede i en ordbogssammenhæng - er den såkaldte *aristoteliske definition*, også kaldet *definitionen via nærmeste genusbegreb (overbegreb) og adskillende kendemærke (differentia)*, eller som definitionsformlen lød i skolastikken: "definitio per genus proximum et differentiam specificam". Formlen kan udtrykkes således:

Definiendum = Definiens:  $g + a_1 + a_2 + \dots + a_n$

hvor "definiendum" er betegnelsen for det udtryk, der skal defineres, og "definiens" er betegnelsen for det udtryk, definiendum skal defineres gennem. I eksplicationen af definiens' form står  $g$  for nærmeste genusterm<sup>3</sup> og  $a_i$  for et vilkårligt adskillende kendemærke. Et lidt barokt klingende eksempel på en sådan definitionstype, hentet hos Aristoteles selv, er:

menneske = dyr + tobenet + fjerløst

eller skrevet i normalprosa: "mennesket er et tobenet og fjerløst dyr". Her er 'dyr' genusbegrebet og 'tobenet' og 'fjerløst' de kendetegn ved begrebet 'menneske', der gør, at det adskiller sig fra begreber for andre dyr. Den ovenfor anførte definition fra *Svensk ordbok* har også denne aristoteliske form med 'anskuelse' som nærmeste genusbegreb og 'radikal' som adskillende kendemærke. Derimod kan "marinere" ikke umiddelbart underlægges dette definitionsskema.

Et specielt karakteristikon ved den aristoteliske definition er *genusbegrebet*. Genusbegrebet er kendetegnet ved, at det skal være mere abstrakt end det begreb, definiens står for. Det kan også formuleres som, at det altid skal være muligt at slutte fra definiensbegreb til overbegreb, men aldrig omvendt:

menneske => dyr                    \* dyr => menneske

radikalisme => anskuelse        \* anskuelse => radikalisme.

(dvs.: alle mennesker er dyr, men ikke alle dyr er mennesker; eller: radikalismen er en anskuelse, men ikke alle anskuelser kan karakteriseres som radikalisme).

Man kan let se, at tanken om definition ved genusbegreb vil kunne videreføres fra definition til definition. Fx vil vi, når vi kommer til definitionen af ordet "dyr", altså begrebet 'menneskes' genusbegreb, igen kunne spørge til *dets* overbegreb (der fx kunne tænkes at være 'levende væsen'), og således videre, indtil vi til sidst vil kunne se alle begreber i sproget organiseret som *kæder af indbyrdes hierarkiserede begreber*.

I et sådant begrebssystem kommer også de adskillende kendemærker til at spille en rolle. De må nemlig fordeles hen over de enkelte sidebegreber på en sådan måde, at ethvert begreb må adskille sig fra ethvert andet begreb med mindst et adskillende kendetegn. Det er nemlig betingelsen for, at de overhovedet vil kunne adskille sig fra hinanden. Dette princip, der ligeledes skal gentage sig ned gennem systemet, kaldes undertiden princippet om inkompatibilitet (termen er lånt fra Lyons, jf. Lyons 1977, 288-289).

På den måde vil vi med udgangspunkt i den aristoteliske definitions bestemmelser være nået frem til en bestemmelse af sproget *som i virkeligheden et eneste omfattende system af begreber og genusbegreber*. Systemets struktur er så enkel og overskuelig, at den kan beskrives simpelt matematisk, nemlig som en semilattice.

Et sådan begrebsstruktur er den gennemgående tanke i videnskabelige klassifikations-systemer som Linné's botaniske taksonomi, ligesom det er ideen bag seksten- og syttenhundredtallets forestilling om et universelt sprog, der i mere praktisk udformning stadig kan spores i forskellige fagtesaurusser eller i almensproglige tesaurusser som fx Roget's Thesaurus fra 1854 (jf. Sparck Jones (1986), 201-203).

Samme tanke finder vi i moderne lingvistik, nemlig som antagelsen om, at det naturlige sprogs leksikon-komponent kan opfattes som et *omfattende system af distinktive træk* (Katz og Fodor 1963).

Endvidere findes tanken videreført i *modelteoretisk semantik*. Ganske vist gør man her op med det ueksplicerede meningsbegreb hos Katz og Fodor, idet man yderligere bestemmer mening som tilordning af tegn i en syntaktisk kalkule - fx en prædikatskalkule (af første orden) - til en semantisk model, hvilket også kan udtrykkes ved, at man går fra en intensionel til en ekstensionel eller referencesemantisk definition af mening (jf. de tidligere intensionelle

og ekstenionelle definitioner af synonymi); men mening forstås stadig som underlagt principperne om synonymitet og meningskompositionalitet.

#### 4. Definition og stereotypi

Der er i de seneste år fremkommet en del kritik af denne måde at forstå og analysere sprog på. Blandt andet er kritikken rejst af Hilary Putnam, hvis teori om stereotyper i bl.a. Putnam (1972) og Putnam (1975) jeg vil belyse i dette afsnit.

Kritikken har haft sine forløbere. Først og fremmest Quine (1951), der endog fremfører kritikken i langt mere radikal form end Putnam. Jeg vil først omtale Quines kritik for derefter straks at vende tilbage til Putnam's mere modererede version.

Det er Quines hovedudgangspunkt, at begrebet 'synonymitet' (og det dermed beslægtede begreb 'analyticitet', der også omfatter andre betydningsrelationer som fx hyponymi) ikke kan forklares cirkelfrit. I konsekvens heraf forkaster han begrebet, idet han opfatter det som et produkt af en fejlagtig teori om forholdet mellem sprog og virkelighed: i den klassiske teori er meningsindhold klart afgrænselige fra hinanden, og det er altid muligt at holde dem ude fra empiriske forhold; en sætning som "en ungkarl er en ugift mand" kan ifølge den klassiske teori vises at være sand *uanset, hvad der empirisk er tilfældet*, mens en sætning som "Jens er ungkarl" omvendt altid kan vises at være empirisk, dvs. oplysende om kendsgerninger i verden. Men det er for Quine en illusion: såvel i dagligdag som i videnskab foregår videnstilegnelse ved, at vi danner os hypoteser (og mere omfattende antagelser), som vi til stadighed tester og vurderer i lyset af de erfaringer, vi indhøster. Men for at forklare dette, er der for Quine at se ingen brug for nogen antagelse om synonymi. Antagelsen om, at der skulle findes hele to mekanismer, én, hvor vi reviderer vores opfattelse af tingene i verden, mens vi holder betydningen af vores sproglige udtryk konstant, og én, hvor vi omvendt reviderer vores sproglige udtryks mening (evt. ud fra pragmatiske kriterier og i lyset af empiriske konstateringer), er for Quine i virkeligheden en kimære.

Men det betyder igen, at afgørelser om synonymitet - og dermed om definitioners pålidelighed - ikke er noget, vi blot undertiden og i visse randtilfælde kan tage fejl i, sådan som det kan tillades i de semantiske systemers opbygningsfase inden for den klassiske opfattelse af definitionen; vanskeligheden er af *principiel* art (med mindre vi ad hoc stipulerer, at det skal være således, hvilket dog gør synonymien pointeløs). Men derfor er også selve begrebet om synonymi ifølge Quine meningsløst, og med det de begreber, der hænger sammen med det, såsom 'mening', 'meningsregel', 'konvention', 'analyticitet', osv.

Nu er Putnam for så vidt enig i størsteparten af Quines kritik: det er korrekt, at vi principielt mangler mulighed for at afgøre, om to udtryk er synonyme. Alligevel er der ifølge Putnam ingen grund til at drage den konklusion, Quine drager, nemlig at det så må indebære, at selve begreberne om synonymi, analyticitet, osv. er meningsløse. Ganske vist har den klassiske definitions tilhængere en for *objektivistisk* forestilling om sproget: det må fastholdes, at sproget ikke er nogen fast afgrænset meningssammenhæng stående som én stor monolitisk blok over for den empiriske verden.<sup>4</sup> På den anden side ender Quines *meningsskepticisme* i det lige så absurde standpunkt, at mening så slet ikke findes, hvilket forekommer ganske kontraintuitivt. Derfor må ifølge Putnam også meningsskepticismen kritiseres. *For faktisk lykkes det os jo normalt udmærket at forstå hinanden* (også i ikke-situationsbundne kontekster). Derfor kan meningsbegrebet ikke blot affejes, men må i stedet omdefinieres. Og det er så det, Putnam forsøger at gøre:



How do I convey the meaning of the word 'tiger'. *I tell him what a tiger is.*  
(Putnam 1981, 147)

Det er i sin korthed Putnams recept. Hermed mener han, at vi skal se langt mere *praktisk* på definitionen i stedet for at lade os blænde af et objektivistisk ideal - eller tomt negere det som Quine: der eksisterer nemlig *praktiske acceptabilitetskriterier* til afgørelse af, om en person har forstået betydningen af et ord eller ej:

The hearer ... already shares a great deal of language with you; otherwise you could'nt tell him what a tiger is. When you "tell him what a tiger is", you simply tell him certain sentences that you accept." (Putnam 1981, 147)

Men hvad er disse sætninger?

The answer runs as follows: there is somehow associated with the word 'tiger' a *theory*; not the actual theory we believe about tigers, which is very complex, but an oversimplified theory which describes a, so to speak, tiger *stereotype*. It describes in the language we used earlier, a *normal member* of the natural kind. (Putnam 1981, 148)

Dette kan man naturligvis bygge "teorier" over. Imidlertid vil disse "teorier" iflg. Putnam have en udelukkende praktisk og anbefalende orientering. At afgøre, hvad definitionen af ordet "æble" er, kræver ingen metodisk anstrengelse, men kan finde sted i en ganske afslappet atmosfære: "hvad forstår du ved et æble?" - "Jo, det er en rund frugt, vi ofte spiser, som er sød og saftig, og som har en ganske tynd grøn, gul eller rødlig skal." - "Aha".

Det betyder *ikke*, at alle æbler så har disse egenskaber, eller at ingen andre frugter så vil kunne have disse egenskaber. Det vil netop være at anstrenge synonymibegrebet i objektivistisk retning. Det betyder blot, at sådan ser æbler nu engang *stereotypisk* ud for os. Det er det billede af et æble, der foresvæver os, i vor daglige kommunikation med hinanden:

In ordinary parlance a 'stereotype' is a conventional (frequently malicious) idea (which may be wildly inaccurate) of what an X looks like or acts like or is. (Putnam 1981, 249)

(Som det fremgår af de to definitioner, er der desværre en uheldig tvetydighed i Putnams begreb om stereotypier: en stereotypi er både noget ekstensionelt - en bestemt, evt. forestillet, genstand - og noget intensionelt - begrebet om, hvad denne genstand er. For at undgå misforståelser vil jeg i det følgende lægge mig op ad definition 2 og reservere betegnelsen "stereotypi" for begrebet eller meningen, mens jeg - med allusion til nyere prototypeteori (jf. Lakoff 1987, 12-57) - vil kalde den genstand, forestillingen er rettet mod, for en *prototype*.)

Putnam giver ingen 'operationelle' anvisninger på, hvilke typer af stereotypier, der findes. Det eneste kriterium, han tilbyder, er ovennævnte acceptabilitetskriterier. Det har givet anledning til en diskussion af, om stereotypiteorien angår alle ord, eller blot de ord, Putnam selv giver eksempler på, nemlig ord for naturlige arter (som tiger, guld, osv.). Selv hælder jeg til den anskuelse, at teorien - især hvis man ser den i lyset af Putnams teori om 'linguistic division of labour' - stort set gælder globalt: for naturlige arter, naturlige egenskaber, artefakter, institutioner, osv.

Det betyder dog ikke, at der ikke kan vises visse tendenser. Heriblandt tendensen blandt folk til at foretrække beskrivelser, der angår forhold tæt på vores sanser og i vores umiddelbare handlesfære, og som ikke er for abstrakte eller specialiserede<sup>5</sup>, sådan som bl.a. Rosch (Rosch 1977) og Lakoff (Lakoff 1987) har peget på: et æble er fx først og fremmest en æble og en frugt. Ikke et Ingrid Marie eller et Cox Orange (det er for konkret). Ejheller en madvare, eller et frøgemme (det er for abstrakt).

Det er klart, at denne nye opfattelse af mening som stereotypi har konsekvenser for den klassiske definition. Det klassiske meningsbegreb falder ganske vist ikke totalt sammen, sådan som det ville være konsekvensen af Quines kritik. Men det må justeres.

For det første har Putnams teori konsekvenser for det indbyggede princip om synonymi. Princippet kan ikke længere opretholdes i sin fulde klassiske udstrækning, hvor der jo generelt skal eksistere sammenfald mellem intension og ekstension (jf. definitionerne af synonymi ovenfor): når der ikke længere findes klare kriterier for kategorimedlemskab, mister kravet om, at to ord i alle kontekster skulle kunne udskiftes med hinanden, overhovedet sin mening. Hvis princippet skulle opretholdes, ville det fx indebære, at "æble" skulle kunne være synonymt med "pære" (da nemlig visse pærer ville kunne tilfredsstille det kriterium for kategorimedlemskab, der er nedlagt i den stereotypiske beskrivelse af "æble").

Imidlertid kan man drage en vigtig lære af Putnams stereotypibegreb: forankres synonymibegrebet *alene* i stereotypibegrebet, ser man faktisk ud til stadig at kunne opretholde princippet. Dog nu kun inden for rammerne af *et helt nyt meningsbegreb, hvor alene prototyperne vil være garanter for sammenfaldet mellem intention og ekstension*: når et æble er "en rund frugt, vi ofte spiser, og som er sød og saftig og har en ganske tynd grøn, gul eller rødlig skal", så gælder det kun for det prototypiske æble. Ikke for alle æbler.

Dette får dog igen konsekvenser for overholdelsen af princippet om meningskompositionalitet. Hvad den kombinerede anvendelse af synonymiprincippet og kompositionalitetsprincippet (plus inkompatibilitetsprincippet) angår, sker der nemlig følgende: når synonymiprincippet forankres i prototypen, bliver synonymi i højere grad bundet til "lokale" forhold: opmærksomheden rettes mod specielt prototypen, ikke så meget mod enkelt-eksemplarenes fællestræk og deres afgrænselighed over for hinanden. Det betyder fx, at æbler kan finde slægtskab med pærer og blomster (de er alle frugter), men vanskeligt med tomater og agurker (der jo er grøntsager), og formentlig slet ikke med bog og agern (der jo slet ikke spises og som i øvrigt kun findes i skoven). Modsætningen til den videnskabelige taksonomi, der forsøger at leve op til de klassiske idealer, er slående. Kompositionalitet *kan* ganske vist findes i almensproglige sammenhænge. Men kun i små miniature-systemer. Længere ude bliver relationerne uklare. Derfor finder vi ikke længere den klare idealstruktur, objektivisternes drømmer om, men kun, hvad vi med et begreb fra Wittgenstein (1953) kan betegne som *familielighed*: drengen ligner godt nok sin broder på næsen, der igen ligner sin søster på hagen, og sin far på øjnene. Men der er ikke længere noget gennemgående træk, som man kan identificere alle familiemedlemmerne gennem.

For det klassiske definitionsbegrebs vedkommende må vi altså konkludere, at i modsætning til, hvad der er konsekvensen af Quines meningsskepticisme, tilkender Putnam med sin stereotypiteori det klassiske definitionsbegreb en vis plausibilitet. Begrebet tåler dog ikke længere den objektivistiske anstrengelse, det tidligere har været udsat for: vi må give slip dels på forestillingen om klart klassememberskab, dels på forestillingen om sproget som ét stort, afklaret begrebssystem.

Spørgsmålet er, om dette valg kan betyde noget for den traditionelle definitionspraksis i ordbøger.

## 5. De to regler for fornuftig definitionspraksis i ordbøger

Her må man vel sige: både - og. For på den ene side er Putnams tanker på mange måder i tråd med fremherskende opfattelser inden for filologien, ikke mindst sådan som den har været præget af tysk romantik: sproget er en mangfoldigt sammensat organisme, ikke noget overskueligt system. På den anden side har der i praksis alligevel været tendenser til at tilstræbe en eller anden form for videnskabelig nøjagtighed i ordbogsdefinitionen, som så til gengæld også - når det i enkelte tilfælde er gået til excesser - er blevet fremholdt til skræk og advarsel. Så alene af den grund synes Putnams betragtninger at kunne være til nytte. Som en art advarsel mod excesser.

Men man kan få endnu mere nytte af hans betragtninger end som så. Man synes nemlig at kunne opstille en art *økonomiprincip for definitioner*, som vil kunne virke styrende ind på udformningen af dem. Princippet, som jeg vil lade være den første af mine regler, lyder:

**REGEL 1:** *Præciser aldrig en stereotypi gennem tilføjelse af yderligere adskillende kendemærker (dvs. beskrivende elementer). Læg hellere - hvis der er plads til det, og hvis det af andre grunde forekommer rimeligt - endnu en stereotypi til den stereotypi, du i forvejen har valgt at bruge i definitionen.*

Som man kan se, skal reglen virke som en slags afbalanceringsprincip, hvor afbalanceringsprincippet yderligere kan skærpes gennem tillægskriterier, dikteret af ordmaterialet: er der tale om et fagord, der har været gjort til genstand for en mere eller mindre eksplicit præcisering, sådan som det er tilfældet i fx fagordbøger og -tesaurusser, er beskrivende elementer selvfølgelig nødvendige. Er der derimod tale om et almensprogligt ord, må beskrivende elementer opfattes som stort set overflødige. Stereotypien er tilstrækkelig.

Regel (1) er først og fremmest en regel, der skal hindre "videnskabeliggørende" præciseringer af definitioner. Forsøg på "videnskabeliggørelse" findes legio i fx *Ordbog over det danske Sprog*, hvor vi blandt andet kan finde følgende forsøg: *dør* = "en til ind- og udgang beregnet aabning i væggen af et lukket opholdsrum (værelse, bolig, hus el.lign.), i alm. forsynet med bevægeligt lukke." (betydn. 1). Her skulle det ikke være svært at finde overflødige præciseringer: *døren* er en ind- og udgang, den er en åbning i væggen til et lukket rum, der er et opholdsrum, den er i alm. forsynet med bevægeligt lukke osv., osv.

Men regel (1) finder ikke kun anvendelse i sådanne klare tilfælde. Også hvor definitionerne ser mere uskyldige ud, kan reglen med fordel anvendes. Fx vil den uden videre kunne hindre definitioner af typen: *hvid* = "som har färg som *nyfallen snö*" (*Svensk ordbok*), *grøn* = "som har en farve som ligner fargen på *friskt gras*" (*Bokmålsordboka*). Her er "nyfallen" og "friskt" ud fra teorien ganske overflødige præciseringer: ifølge vor stereotypier om de respektive arter er sne altid hvid og græs altid grønt. De behøver ingen præcisering for udelukkelse af specialtilfælde.

Regel 1 er som sagt udtryk for et afbalanceringsprincip. Det betyder ikke, at der så ikke må findes præciseringer overhovedet, men det indebærer, at man i stedet for præciseringer - hvor det er rimeligt - hellere skal finde en eller flere ekstra stereotypier. Eksempelvis, at man i definitionen af "hvid" i stedet for "har farve som nyfalden sne" hellere skal foretrække fx "har farve som sne og sukker" eller "har farve som sne og mel".

Men hvornår er forholdene rimelige? Det hænger i høj grad sammen med, hvad definitionen skal bruges til.

Undertiden er der tale om, at en ordbogsbruger ikke kender det begreb, der knytter sig til et ord: i sådanne tilfælde må et større sæt af stereotypier vælges. Jo flere, der nævnes, jo mere greb vil ordbogsbrugeren - der kan være en skoleelev, en person fra et ikke-bogligt orienteret socialt lag eller en indvandrer - få om ordet. Eksempelvis er der større chance for, at en indvandrer fra en anden kulturkreds vil lære betydningen af ordet "hvid", hvis der står *både* "sne" og "mel". Han kender måske slet ikke til sne, eller har kun en vag forestilling om, hvad sne er, eller han kender måske til sne, men har aldrig før set hvidt mel.

I de fleste tilfælde, er der dog tale om, at ordbogsbrugeren på forhånd kender til begrebet. Det er blot det tilsvarende ordudtryk, der mangler. Det vil fx gælde for den udlænding, der tilhører en nærtstående kulturkreds, fx en moderne bykultur. I sådanne tilfælde er det normalt ikke nødvendigt, at brugeren præsenteres for et mere omfattende sæt af stereotypier. Her er det tilstrækkeligt med det sæt af stereotypier, (på måske blot én stereotypi), der gør ham eller hende i stand til at *aktivere* resten af sættet.

Men derfor kan vi formulere følgende regel i tillæg til regel 1:

**REGEL 2:** *Vælg i en definition ud over din genusterm netop så mange stereotypier som differentia, at det er tilstrækkeligt til at aktivere ordbogsbrugeren eget, implicite beredskab af stereotypier. I det omfang du også vil beskrive sættet af stereotypier, er det kun rimeligt i det omfang, ordbogsbrugeren ikke i forvejen kender til stereotypien.*

Nu er der selvfølgelig mange mellemløsninger mellem præsentationen af blot én stereotypi i en definition og så et mere omfattende sæt af stereotypier. Det rejser blandt andet spørgsmålet om, hvor mange stereotypier, der findes til et givet ord (ud over det eller de stereotypier, der er tilstrækkelige til at identificere ordbetydningen).

For så vidt er spørgsmålet meningsløst. Antallet af stereotypier må afhænge af antallet af mulige praktiske anvendelsessammenhænge for ordet. Men dette antal ligger selvfølgelig ikke fast. Det er klart, at en marsboer (eller et ekspertsystem) må have flere stereotypiske oplysninger om et ord, end en indfødt sprogbruger, der blot lige netop ikke kender dette ord. Men det er svært i almindelighed at sætte tal på stereotypierne knyttet til et bestemt ord. Wierzbicka (Wierzbicka 1985) har specificeret sæt af stereotypier for forskellige ord. De fylder ca. 1 - 2 tekstsider hver.

Et aspekt af denne problematik er selvfølgelig opløsningen af et ord i polysemer: når man i *Collins Concise Dictionary and Thesaurus* definerer "vej" som gade (road = street), kan man måske synes, at den stereotypi, definitionen skal trække på, er lidt vel langt væk: "vej" kan måske nok eksemplificeres, men vanskeligt *defineres* ved gade. Men igen må det langt hen ad vejen være et spørgsmål om ordbogens totale tilrettelæggelse og hensynet til brugergruppen.

Man kan operere med en niveaudeling af definitionerne: *vej* 1. det sted, hvor biler og cykler kører ... 1.1. landevej ... 1.2. vej i by; gade. (denne præsentationsmåde løber risikoen for pseudopræcisering: det kan være, det slet ikke er rimeligt at påtvinge stereotypierne denne struktur). Man kan vælge at operere med en sideordning: *vej* 1. landevej ... 2. vej i by; gade (denne sideordning er *Cobuild's* overvejende strategi, selv om det ikke lige gælder ved dette ord). Eller man kan vælge at operere med kun én betydning, enten præsenteret ved en mere eller mindre central stereotypi - *vej* det sted, hvor biler og cykler kører - eller ved to mere eksemplificerende stereotypier - *vej* landevej, gade.

Imidlertid synes man ikke på forhånd at kunne opstille retningslinier for, hvilken præsentationsmåde, der er den bedste. En række forhold vil kunne gøre sig gældende.

Som konklusion på diskussionen i det foregående må man sige, at selv om Putnams stereotypeteori indebærer et opgør med det klassiske definitionsbegreb, er der på ingen måde tale om nogen forkastelse af begrebet. Tværtimod står ideen om definition ved genusbegreb og adskillende kendemærke i en vis forstand styrket. Ganske vist kræver det så til gengæld, at man tager hensyn til de præciseringer, der kan udledes af Putnam's teori, og - ikke mindst - at man afstår fra tilbøjeligheden til at pseudopræcisere og pseudovidenskabeliggøre.

Det skulle mine to regler gerne sikre.

## 6. Afslutning

Tilbage står stadig et stort arbejde med at udvikle teorien - og mine regler - yderligere. Her kan overvejelserne inden for cognitive science (frame-teori, script-teori, teorien om idealized cognitive models, prototypeteori og hvad det ellers kan hedde) formentlig bidrage med interessante resultater til gavn for definitionsarbejdet.

Et spørgsmål, jeg i den forbindelse ikke har berørt i artiklen, men som bl.a. synes at have betydning for genusbegreberne, er spørgsmålet om, hvorvidt det er muligt at etablere en fornuftig *ordning* blandt stereotypierne<sup>6</sup>.

At rejse dette spørgsmål, er samtidig at rejse spørgsmålet om, hvorvidt det alligevel skulle vise sig muligt at finde en *overordnet* semantisk systematik - i det mindste for visse lag i sproget - til trods for den hårde kritik af de overdrevne systematiseringsbestræbelser inden for det klassiske definitionsbegreb.

Putnam's stereotypeteori udgør blot en lille del af hans samlede sprogteori. Den er indlejret i den mere omfattende teori om 'linguistic division of labour'. I andre dele af denne teori mener Putnam faktisk at kunne give en (transcendental) begrundelse af mere overordnede eller centrale lag i sproget. Og hvad der forekommer ganske interessant set ud fra en systematisk synsvinkel: han mener også herigennem at kunne forsvare en "nominel" anvendelse af den modelteoretiske semantik.

Om Putnam har ret, og om det vil lykkes at "indpode" en sådan systematik, må fremtiden vise. Mange forsøg herpå er i gang. Jeg har bl.a. selv tidligere søgt at skitsere en sådan systematik (Widell 1989 og Widell 1990).

Hvorom alting er, må man gøre sig en ting klart: det er ikke længere muligt at opfatte sproget *som helhed* som en sådan struktur. Denne mulighed er - efter Quine's og Putnam's kritik - endegyldigt udelukket. Derfor vil det kun være gennem idealiserende abstraktioner det vil være muligt at tegne strukturelle totalvuer.

Et spørgsmål, der stadig står hen - set ud fra en leksikologisk synsvinkel - er, om abstraktionerne og redegørelserne her vil falde så komplicerede og esoteriske ud, at strukturanalysen i lingvistikken og stereotypi-belysningen i leksikologien i virkeligheden vil vokse fra hinanden. Da må ønsket om en multifunktionel base opgives. Vi ender da måske til slut i den situation, at vi vil have "Lexicons for Computer Programs" på den ene side og "Lexicons for People" på den anden side (jf. titlen på en artikel af Kay, Kay 1989).

## Noter

1. Det skal understreges, at opfattelsen af den intensionelle og den ekstentionelle definitions måde som ækvivalente definitions måder ikke er uproblematisk. Den bevirker nemlig, at distinktionen mellem synonymi og ko-referentialitet udviskes. Det har fx været et problem ved Quines meningsskepticisme. Vi vil vende tilbage til Quine senere i artiklen.

2. I ordbogsdefinitionen indgår normalt ud over den semantiske antagelse om synonymi også en pragmatisk antagelse om, at det ord, opslagsordet defineres gennem, skal være bedre kendt end selve opslagsordet: hvis det nemlig var omvendt, ville ordbogsbrugeren ikke få noget ud af ordbogen. Dette forhold vil vi ikke berøre i denne sammenhæng.

3. Hvad der er nærmeste genusbegreb i de konkrete tilfælde, er selvfølgelig ikke angivet med formlen her. Eksemplerne skulle gerne anskueliggøre, hvad der kan menes. Det kan tilføjes, at spørgsmålet synes at kræve et struktureret begrebssystem, der fastlægger afstandene mellem de enkelte begreber. Ellers vil det være vanskeligt at tale om et *nærmeste* overbegreb. Da det netop er et hovedanliggende at problematisere en universel anvendelse af begrebssystemer med genusbegreber, vil jeg ikke berøre forholdet nærmere.

4. Det er vigtigt, man får øje på, at objektivismen ikke er nogen esoterisk opfattelse, som blot kan tilbagevises ved, at man indtager en mere moderat holdning. Det er ikke tilstrækkeligt, at man siger: "Jeg vil ikke stille sproget *in extenso* over for verden, for det virker - korrekt nok - forrykt. Jeg vil nøjes med at stille blot en lille del af sproget over for verden. Så synes der at kunne komme fornuft i det, *for så har vi med noget at gøre, som vi kan overskue*. Resten kan så defineres på anden vis, fx gennem synonymi." En sådan holdning er ikke mulig. For indtager man den over for en del af sproget, vil man *nødvendigvis* indtage den over for sproget som helhed. Det springende punkt er *eksternalismen*: at man begynder at se forholdet mellem sprog og verden som et ydre forhold mellem to selvstændige, på forhånd afgrænsede, komponenter (ord - ting, sætning - sagforhold, sprog - verden).

5. Lakoff kalder det første træk *conceptual embodiment*, det andet - følgende Rosch - *basic level categorization* (Lakoff 1987, 12-13). En række andre træk ved prototypisk definerede begreber ud over disse to nævnes også af Lakoff, men det vil føre for vidt i denne sammenhæng at introducere dem.

6. Wierzbicka har i Wierzbicka 1985 været inde på spørgsmålet. Ligeledes Lakoff i Lakoff 1987.

## Litteratur

- Byrd, Roy J. & Calzolari, Nicoletta & Chodorow, Martin S. & Klavans, Judith L. & Neff, Mary S. & Rizk, Omneya A. 1987. Tools and Methods in Computational Linguistics. *Computational Linguistics* (1987), Volume 13, 3-4, 219-240.
- Boguraev, Bran & Briscoe, Ted, red. 1989. *Computational Lexicography for Natural Language Processing*. London & New York.
- Calzolari, Nicoletta. 1990. Structure and Access in an Automated Lexicon and Related Issues. I: Donald Walker, Antonio Zampolli & Nicoletta Calzolari, red. *Automating the Lexicon: Research and Practice in a Multilingual Environment*.
- Eikmeier, Hans-Jürgen & Rieser, Hannes, red. 1983. *Words, Worlds, and Contexts: New Approaches in Word Semantics*. Berlin, New York.
- Fodor, Jerry A. & Katz, Jerold J. 1963. The Structure of Semantic Theory. I: *Language*, vol. 39. 170-210.
- Harras, Gisela & Haß, Ulrike & Strauß, Gerhard. 1991. *Wortbedeutungen und ihre Darstellung im Wörterbuch*. Berlin, New York.

- Hudson, Richard. 1988. The Linguistic Foundation for Lexical Research and Dictionary Design. *International Journal of Lexicography*, Vol. 1, No. 4. 287-312.
- Jones, Karen Sparck. 1986. *Synonymy and Semantic Classification*. Edinburgh.
- Kay, Martin. 1989. Lexicons for Computer Programs and Lexicons for People. I: *Dictionaries in the Electronic Age, Fifth Annual Conference of the UW Centre for the New Oxford English Dictionary - Proceedings of the Conference*. 43-67. Oxford.
- Lakoff, George. 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago.
- Lyons, John. 1977. *Semantics 1*. Cambridge.
- Putnam, Hilary. 1981. *Reason, Truth and History*. Cambridge.
- Putnam, Hilary. 1975. Is semantics possible? I: *Mind, Language and Reality: Philosophical Papers, volume 2*. 139-152. Cambridge.
- Putnam, Hilary. 1975. The meaning of 'meaning'. I: *Mind, Language and Reality: Philosophical Papers, volume 2*. 215-271. Cambridge.
- Putnam, Hilary. 1983. Reference and truth. I: *Realism and Reason: Philosophical Papers, volume 3*. 69-86. Cambridge.
- Quine, W. V. 1951. Two dogmas of empiricism. I: *From a Logical Point of View*. 20-46. Cambridge, Massachusetts.
- Rosch, E. 1977. Human Categorization. I: N. Warren, red., *Studies in Cross-Cultural Psychology*. London.
- Sinclair, J. red. 1987. *Looking Up*. London.
- Stein, Gabrielle. 1979. The best of British and American lexicography. I: *Dictionaries 1* (1979). 1-23.
- Walker, Donald E. 1989. Developing Lexical Resources. I: *Dictionaries in the Electronic Age, Fifth Annual Conference of the UW Centre for the New Oxford English Dictionary - Proceedings of the Conference*. 1 - 22. Oxford.
- Widell, Peter. 1989. A Model for Constructing the Basic Categories in a Natural Language Thesaurus. I: Karl Hyldgaard-Jensen & Arne Zettersten, red. *Proceedings of the Fourth International Symposium on Lexicography April 20-22, 1988 at the University of Copenhagen*. Tübingen. 33-45.
- Widell, Peter. 1990. THESPART - en multifunktionel semantisk database. I: *Årsberetning fra Selskab for Nordisk Filologi*. København
- Wierzbicka, Anna. 1985. *Lexicography and Conceptual Analysis*. Ann Arbor
- Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford. (Dansk oversættelse: *Filosofiske Undersøgelser*, København 1971).

#### Ordbøger:

- Collins Cobuild English Language Dictionary*. 1987. London, Glasgow.
- Collins Concise Dictionary and Thesaurus*. 1991. Glasgow.
- Nudansk Ordbog*. 1990. København.
- Ordbog over det danske Sprog 1-28. 1918-1956*. København.
- Webster's Seventh New Collegiate Dictionary*. 1963. G. & C, Merriam. Massachusetts.

## Tvärspråklig lexikologi med svenskan i centrum<sup>1</sup>

Med tvärspråklig lexikologi menas en beskrivning av ordförrådet som baserar sig på en jämförelse mellan olika språk. I den här artikeln kommer jag först att ge en kort generell översikt av ämnesområdet för att sedan kort karakterisera mina egna lexikala studier av lexikala universalier, svenskans lexikala typologiska profil och lexikal inläring.

### 1. Vad är tvärspråklig lexikologi?

Språk kan jämföras i många syften. Den tvärspråkliga (eng. *crosslinguistic*) jämförelsen tar främst fasta på universella och typologiska drag. *Universella drag* utgörs av sådana egenskaper som delas av alla språk eller av generella restriktioner för vad som är möjligt i ett språk medan *typologiska drag* utgörs av egenskaper som visserligen växlar mellan språken men är så centrala att de utsäger något om den grundläggande strukturen inom något område (t ex verbordförrådet, ett grundläggande semantiskt fält). Egenskaper som är mer eller mindre unika för ett språk eller (oftast) en grupp genetiskt och/eller arealt nära relaterade språk kallas *språkspecifika*. Även inom ett språk uppvisar språkspecifika drag ofta mer komplexa egenskaper i olika avseenden än universella drag, vilket är utgångspunkten för *markeringsteorin* (Greenberg 1966, Croft 1990). Denna teori tar fasta på att vissa drag, som kallas *omarkerade*, intuitivt uppfattas som mer grundläggande, naturliga eller enkla. Omarkerade drag förekommer i fler språk än markerade samtidigt som de t ex har högre frekvens i enskilda språk och representerar en föredragen tolkning i fall av neutralisation. Sådana likheter mellan språk som återspeglar generella struktureringsprinciper måste skiljas från överensstämmelser som beror på historisk släktskap (*genetiska drag*) och på språkkontakt som ofta yttrar sig som en bestämd geografisk spridning (*areala drag*). Teoretiskt är syftet med den tvärspråkliga jämförelsen att undersöka vad som är ett möjligt mänskligt språk och att kunna förklara hur den språkliga, i detta fall lexikala, organisationen interagerar med biologiska, kognitiva och sociokulturella faktorer. Praktiskt är syftet detsamma som för den kontrastiva analysen, nämligen att ge ett förbättrat underlag för språkundervisning, översättning och tolkning. Skillnaden är att den kontrastiva analysen ofta inskränks till en jämförelse mellan par av språk utan att man kan avgöra vilket av språken som eventuellt uppvisar ett markerat alternativ.

---

<sup>1</sup> Arbetet med denna uppsats har utförts inom min tjänst i lexikalisk forskning som finansieras av Humanistisk-samhällsvetenskapliga forskningsrådet.



## 2. Lexikala universalier och svenskans lexikala profil

Jag ska inte här försöka ge någon allsidig översikt över studiet av lexikala universalier utan koncentrera mig på några centrala utgångspunkter för mina egna studier som rör grundstrukturen hos de öppna ordklasserna substantiv, adjektiv och verb med den sista ordklassen i fokus. Orden inom dessa ordklasser kan indelas i semantiska fält. De mest grundläggande av dessa realiseras i alla språk, där ordklassen föreligger. Förbehållet gäller för adjektiven. Dixon (1977) har visat att om ett språk har adjektiv finns alltid fält som Dimension (stor/liten), Ålder (gammal/ung), Värdering (bra/dålig) och Färg (svart/vit). Det finns vissa språk där endast ett litet antal ord med de angivna betydelseerna realiseras som adjektiv. För färgordens del har Berlin & Kay (1969) visat att det finns en lexikaliseringshierarki, där de första stegen utgörs av svart/vit > röd > blå, grön etc. Det går således att förutsäga hur fältet byggs ut. I europeiska språk som har tämligen omfattande klasser av adjektiv återspeglas den omarkerade naturen hos dessa adjektivbetydelser genom att adjektiv med motsvarande betydelser hör till de mest frekventa i språket (Viberg 1990).

För verbens del gäller att det även i detta fall finns vissa språk med en mycket begränsad (12-40) uppsättning av enkla verb (medan frasverb av typen kasta ett öga på är mycket talrika). I alla dessa språk finns dock vissa *kärnverb* (alternativt: *nukleära verb*) med en karakteristisk grundbetydelse som tämligen systematiskt fördelar sig på de mest grundläggande verbala semantiska fälten: 'gå' (fält: Förflyttning), 'ge' (Ägande), 'slå' (Fysisk kontakt), 'säga' (Verbal kommunikation) och 'se' (Perception). Europeiska språk har i allmänhet ett stort antal verb (vanligen i storleksordningen 10 000). Trots detta råder en stark koncentration till några få mycket vanliga verb i normal diskurs (text). De tjugo vanligaste verben tenderar att täcka drygt 40% av de löpande verben i stora korpora (på 1 miljon ord eller mer). Vissa av de mest frekventa utgörs av grammatiska verb som kopulan och modala hjälverb, vilka representerar ett arealt specifikt drag i de europeiska språken som ofta saknar motsvarighet i utomeuropeiska språk. Dessutom återfinns bland de tjugo vanligaste verben i ett representativt urval av 11 europeiska språk nästan undantagslöst en liten uppsättning kärnverb som har samma grundläggande betydelse som i språk med minimala verbsystem. (Se Viberg 1990, 1991a för en utförligare beskrivning och motivering.)

Den inre strukturen inom ett semantiskt fält exemplifieras i Viberg (1984a) som behandlar perceptionsverb. På grundval av data från ca 50 språk upprättas en lexikaliseringshierarki som lätt förenklad kan återges som:

'se' > 'höra' > 'känna' > 'förmäta smak, lukt'

Ett verb med betydelsen 'se' återfinns i alla språk som har perceptionsverb markerade för sinnesmodalitet. (Det finns några språk som helt saknar enkla perceptionsverb markerade för en specifik sinnesmodalitet.) Om det finns två verb markerade för sinnesmodalitet utgörs de av 'se' och 'höra' etc. Hierarkin redogör också för möjlig polysemi. I svenskan kan det verb som är knutet till känsel, känna, utsträckas till att täcka även smak och lukt (Kalle kände lukten av nybakade bullar). På samma sätt finns det språk där man kan 'höra' lukt och smak. Men polysemin kan inte gå i motsatt riktning.

Med ett språks *lexikala profil* menas en tvärspråkligt baserad karakteristik av vilka drag som reflekterar universella tendenser eller den lexikala typ som språket tillhör och vilka som är språkspecifika. I Viberg (1990) ges en karakteristik av svenskans grundläggande verbordförråd från dessa utgångspunkter. Bland universella drag uppvisar svenskan de typiska kärnverben såsom *gå*, *ge*, *se*, *säga* etc., som hör till de 20 frekventaste verben i språket. Till

de mest frekventa verben hör också grammatiska verb som kopulan vara, verbet ha och de modala hjälpverben, vilka samtliga representerar arealt specifika drag. Särskilt specifikt är den höga frekvensen för existens verbet *finnas*. Det går också att identifiera exempel på *särskilt semantisk differentiering* och *särskilt polysemimönster*. Exempel av den första typen är serier som *tro/tycka/tänka* (Kognition) och *sätta/ställa/lägga* (Förflyttning) medan exempel av den andra typen utgörs av *få* (Ägande: få ett äpple/ Modal: få stanna), ett polysemimönster som är karakteristiskt för språken i Norden (inklusive finskan). Data från andra språk än svenskan insamlas med en lexikal frågelista (Viberg 1986).

### 3. Fysiska kontaktverb och verbet stryka som exempel på lexikal analys

För närvarande arbetar jag med en studie av svenskans verbordförråd, som består av en studie av polysemin hos de mest grundläggande verben (bl a kärnverben) kombinerad med studier av den semantiska differentieringen inom några valda fält, där jag har ambitionen att täcka alla verb som hör till fältet. Som ett exempel på hur analysen är upplagd har jag valt de fysiska kontaktverben (Viberg 1984b). Dessa har alla kärnkomponenten CONTACT (x, y) 'det råder (uppstår) fysisk kontakt mellan x och y'. Kärnverbet inom fältet är *slå* som är det i särklass mest frekventa och mest polysema verbet inom fältet. Eftersom utrymmet är begränsat, har jag valt ett annat verb nämligen *stryka* som huvudexempel. Ambitionen är nämligen att ge en semantisk karakteristik som täcker alla de exempel som föreligger i representativa konkordanser. Den följande analysen bygger på Språkdats konkordanser över svenska romaner.

Som en typisk användning av fysiska kontaktverb uppfattar jag följande satstyp: Per (Agent) slog/dunkade/petade/strök (Fysiskt kontaktverb) Kalle (Erfarare) i magen/ i ryggen/ i sidan/ över pannan (Mål) med handen/en käpp (Instrument). Många av verben inom fältet inkorporerar en kroppsdel som i det omarkerade fallet inte realiseras på ytan: slå, stryka (hand), slicka (tunga), kyssa (läppar), sparka, trampa, stampa (fot), stånga (horn, panna). De flesta av verben inom fältet inkorporerar också en rörelse och olika riktningar: sparka (röra foten framåt/bakåt), stampa, trampa (upp/ner). Verbet *stryka* hör till en grupp verb som inkorporerar en rörelse längs något (ALONG(TRAVEL)) (x, y): stryka, gnida, gnugga, smeka, kittla m fl.

Följande autentiska exempel med verbet *stryka* innehåller de typiska betydelsekomponenterna. Liksom de övriga exemplen i återstoden av detta avsnitt är det hämtat ur Språkdats romankorpusar.

Baronen strök med handen över ögonen.

Någon (z) utför viljemässigt en handling som innebär att något (x) som i typfallet är agentens hand förflyttas längs något (y) samtidigt som det råder kontakt mellan x och y. Om det senare villkoret inte vore uppfyllt vore det riktigare att säga att baronen förde handen över ögonen. Formaliserat kan grundbetydelsen återges på följande sätt med användandet av det system som beskrivs i Miller & Johnson-Laird (1976), vilka dock inte analyserat fysiska kontaktverb:

- (i) (WITH<sub>i</sub> (ACT)) (z, S, x)  
 (ii) (ALONG(TRAVEL)) (x,y)  
 (iii) HAPPEN (CONTACT (x, y))  
 (iv) CAUSE (S, (ii) & (iii))  
 Default: PALM(x) & POSSESS (z, x)

I det givna exemplet anges x explicit som "handen". Detta är dock snarast ett undantag. Även i följande exempel är emellertid den naturligaste tolkningen att det är handen som förflyttas och kommer i kontakt med objektet:

Grevinnan böjde sig fram och strök Stina över håret

Genom att en person syntaktiskt fungerar som direkt objekt tillkommer i detta exempel en komponent som har paralleller hos många av de fysiska kontaktverben, nämligen en kommunikativ eller expressiv innebörd som är överlagrad den konkreta innebörden. Enligt en generell princip uppfattas det som sker med objektet då detta är specificerat (handlingens resultat) som förgrundad information, varför den expressiva innebörden rycker fram på första plats. För *stryka* kan den expressiva innebörden uttryckas som att z vill visa w (objektet) tillgivenhet:

- (v) INTEND (z, SHOW(z, w, AFFECTION))

I romantexterna framstår strykandet av någon över håret nästan som en schablon för uttryckande av tillgivenhet. I kombination med andra kroppsdelar kan handlingen få en annan, t ex en sexuell symbolisk innebörd, men detta följer av allmänna kulturella regler för de olika kroppsdelarnas symbolvärde. Om exempel med personobjekt uppfattas som de mest prototypiska, kan många exempel på polysemi som finns i materialet ordnas i en skala efter avtagande prototypikalitet, på ett sätt som återges schematiskt nedan:

#### Polysemi: Avtagande prototypikalitet

<i>Exempel:</i>	<i>Aktualiserade komponenter:</i>
Av en låtsad tillfällighet strök han sköterskan över baken.	(i)-(v)
elefantmodern /.../ fortsatte att stryka med snabeln över sitt barns kropp	i) - (iv)
En hög svart vagnskaross strök mot slingerväxterna	(ii) - (iii)
Stormen stryker förbi över havet	(ii)

Det finns visserligen ett betydande antal exempel som faller utanför detta mönster men de kan inordnas i två grupper. Det första som jag inte ska redogöra för här gäller det fall då verbet kombineras med en partikel. På så vis bildas ett frasverb som enligt produktiva regler överför verbet till ett annat semantiskt fält, t ex *stryka av*, *stryka bort*, *stryka ut*

(Avlägsnande/Utplånande). Partikeln anger i detta fall resultatet medan verbet främst anger en sättskomponent som kan kontrastera mot verb hämtade från olika fält (avlägsna, utplåna: stryka bort - skrapa av - plocka bort - spola ren etc). Som helt lexikaliserade måste man betrakta den andra gruppen av exempel, i vilka instrumentet (x) antar olika värden. Liksom handen i det prototypiska fallet behöver instrumentet inte anges explicit. Det finns tre sådana fall:

Jag strök taket med silverfärg. BRUSH (x) & SURFACE(y)

Hon får göra vad hon kan: baka, byka,  
skura golv och stryka kyrkoherdens  
kalsonger. IRON (x) & CLOTH (y)

Han strök under ordet "specialritade" PEN (x) & TEXT (y)

I alla tre fallen förflyttas x längs något. Skillnaden består i det konventionella resultatet som bör anges med en särskild resultativ komponent som fogas till representationen för grundbetydelsen. (Hur detta formaliseras kommer att diskuteras längre fram.) I det övre exemplet förs en färgborste (x) längs taket (y), vilket orsakar att taket förses med färg. I nästa exempel är det ett strykjärn som förs längs ett tyg, vilket leder till att detta blir slätt och i det sista exemplet förs en penna över en text, vilket leder till att det uppkommer ett streck. Det sista exemplet har också en kommunikativ innebörd. Att stryka under markerar att något är viktigt, vilket kontrasterar med att stryka över, vilket markerar att något upphör att gälla. Dessa kommunikativa innebörder bildar dock ett system av kontraster på en semantisk nivå som är överlagrad den rent konkreta och som är helt oberoende av det system av kontraster som råder mellan *stryka* i dess olika användningar som ett kontaktverb och övriga kontaktverb. Schematiskt kan användningen av verbet med 'penna' som instrument beskrivas på följande sätt:

Syntaktisk ram: NP<sup>z</sup> \_\_\_ (för/under/över) NP<sup>y</sup> (för NP<sup>w</sup>) (med NP<sup>x</sup>)  
Selektion: PERSON: z, w; TEXT: y; PEN: x

- (i) (WITH<sub>i</sub> (ACT)) (z, S, x)
- (ii) (ALONG(TRAVEL)) (x,y) & HAPPEN (CONTACT (x,y))
- (iii) CAUSE (S, (ii) )
- (iv:1) CAUSE ( (ii), HAPPEN(EXIST(LINE)))
- (v:1) INTEND (z, SHOW (z,w, { IMPORTANT (y))  
  notVALID (y)

Den kommunikativa innebörden som återges längst ner är beroende av vilken partikel som väljs: *för* och *under* anger att något är viktigt medan *över* anger att texten inte längre är giltig:

Hogarth hade i sin kopia strukit under namnet Tore Hansson varje gång det förekom.

Ibland hade hon strukit över och skrivit "önskedröm".

Även i detta fall kan olika betydelsekomponenter undertryckas. Endast den kommunikativa komponenten finns med, då verbet uppträder med ett abstrakt objekt i uttryck av typen: stryka

under (vikten av) att. Ofta signaleras den abstrakta användningen formellt genom fast sammansättning: understryka. En mer levande metafor fås om man som i följande exempel behåller en hänsyftning på den skriftliga representationen:

Man strök ett streck över åtminstone den slentrianen.

Jag har här endast kunnat skissera en fullt formaliserad beskrivning av verbet *stryka*. Vad som sagts bör dock räcka för att visa hur formalismen i princip gör det möjligt att tydligare visa sambandet mellan olika delbetydelser och de kontextuella villkoren för de olika tolkningarna. En liknande skiss av fälten Diskonnexion, Skada och Reparation ges i Viberg (1985a).

#### 4. Lexikal inläring ur tvärspråkligt perspektiv

Studier av hur ordförrådet lärs in utgör både en praktisk tillämpning och ett test av den psykologiska realismen hos den modell av ordförrådets organisation som skisserats ovan. Kännetecknande för både första- och andraspråksinläringen är att ordförrådet växer långsamt. Att kunna kommunicera är för inläraren av ett språk i stor utsträckning en fråga om att klara sig med få ord. Detta gäller naturligtvis särskilt starkt på tidiga stadier av inläringen. Men problemet kvarstår länge. Inom verbordförrådet finns det en klar tendens att det är ett verb inom varje semantiskt fält som får fylla luckorna inom fältet och detta tenderar att utgöras av ett av de nukleära verben. Till att börja med innebär detta kanske att det får ersätta de flesta andra verben inom fältet. Allteftersom andra verb inom fältet behärskas får verbet fylla de återstående vita fläckarna. Sådana verb utgör en slags kärna kring vilken fältet byggs ut, vilket motiverar benämningen nukleära verb. I Viberg (1991b) redogörs för en studie av hur sexåriga invandrarbarn tillägnar sig svenska som andraspråk. Bandade muntliga samtal jämfördes med liknande inspelningar gjorda med jämnåriga barn med svenska som förstaspråk. De nukleära verben visade sig ha en hög frekvens redan hos de infödda talarna men hade trots detta en stark tendens till överrepresentation hos invandrabarnen. Denna tendens kvarstod då nya inspelningar gjordes ett år senare, vilket tyder på att det starka favoriserandet av ett litet antal nukleära verb är ett drag som förändras långsamt. Samma tendens finns också hos vuxna andraspråksinlärare (Viberg 1988) och hos små barn (2-3 år), då de lär sig sitt förstaspråk (Viberg 1991a).

En viktig skillnad mellan första- och andraspråksinläring tycks vara att språkspecifika (men grundläggande) drag tenderar att dyka upp tidigt vid förstaspråksinläring medan sådana drag har en stark tendens att undvikas eller ge upphov till fel vid andraspråksinläring. På det lexikala planet ger språkspecifik polysemi och differentiering upphov till återkommande problem vid andraspråksinläring. Eftersom inlärarna har en stark tendens att undvika sådana strukturer är det i många fall nödvändigt att tillgripa speciella former av styrd elicitering för att studera inläringen av sådana drag. I Viberg (1985b) redogörs för tillägnet av svenskans placeraverb *sätta/ställa/lägga*. Jag ska här illustrera metoden med en mindre studie av ett annat fall av språkspecifik differentiering i svenskan, nämligen verbparet *skära* och *klippa*. I många andra europeiska språk motsvaras dessa två verb i sin mest typiska användning av ett verb, t ex engelska *cut*, tyska *schneiden*, franska *couper*, nygrekiska *kóvo*. Verben hör till fältet Diskonnexion som omfattar ett stort antal verb organiserade kring betydelsen 'dela i bitar' (bryta, bräcka, spräcka, krossa, skiva etc). I den aktuella studien undersöktes 60 vuxna inlärare av svenska som andraspråk, vilka jämnt fördelade sig på de tre modersmålen, spanska, polska och finska. Tre diskonnexionsverb undersöktes nämligen *skära*,

*klippa* och *riva*. Skillnaderna i differentiering mellan svenskan och modersmålen framgår av figur 1. Som framgår motsvaras *skära* och *klippa* av ett ensamt verb medan *riva* har en motsvarighet, vilken dock i spanskan har en mer allmän betydelse (som eng. break). Alternativt används i spanskan också uttrycket *hacer pedazos* 'göra bitar'.

Semantiskt drag	Språk			
DELA I BITAR	Svenska	Spanska	Polska	Finska
MED EN KNIV	skära	cortar	pociąć	leikata
MED EN SAX	klippa			
GENOM ATT DRA ISÄR	riva	romper	porwać	repiä

(På polska finns egentligen två motsvarigheter till *skära* och *klippa* men de kontrasterar semantiskt på ett annat sätt.)

Figur 1. Diskonnexionsverb i svenskan och källspråken.

Uttryck med de aktuella betydelseerna elicerades muntligt med ett videotest som visade enkla handlingar. Försökspersonerna fick i uppgift att instruera en kamrat i samma undervisningsgrupp som inte kunde se videon att utföra samma handling. I figur 2 visas några representativa exempel på hur betydelsen 'skära' uttrycktes. Det översta alternativet är vad jag själv skulle kunna säga. Inom parentes identifieras inlärarna av ett nummer föregånget av första bokstaven i modersmålet.

<p><i>Idealiserad infödd version: skär av en bit snöre med kniven (ÅV)</i></p> <p>skär liten bit snöre med en kniv (P66)</p> <p>du har en snör / och du har en kniv / och eh / du behöver nu två korta bitar / snö (P1)</p> <p>ta en knif / och / kl // klipp klippa lite liten snö (P2)</p> <p>sen du måste klippa den / med en kniv som / som e röd (F66)</p> <p>och klippa inte inte inte med sax med / kniv (S10)</p> <p>ta skär å skära /.../ den här sak som kvinna estickar (S12)</p>
--

Figur 2. Realisationer av *skära*

Mest slående är att många av informanterna använder *klippa* i stället för *skära*. Åtminstone ett av exemplen (S10) tyder dock på att informanten varit medveten om att *klippa* inte är det korrekta verbet, eftersom hon tillägger "inte med sax med / kniv". Det finns också ett tydligt exempel på en kommunikationsstrategi i form av omskrivning som i "du behöver nu två korta bitar". I figur 3 ges några exempel på hur 'klippa' och 'riva' uttrycktes.

*Idealiserad infödd version:* ta en blå servett å klipp den mitt itu (ÅV)

du måste klippa / ta en / inte kniv men den andra (F64)

ta blåa papper / och / skä skär i två del / me s sax (P68)

du kan vika den och dela / i två delen (P52)

saxa nej / saxa på två delen (P3)

*Idealiserad infödd version:* ta en servett och riv sönder den i små bitar (ÅV)

ta en vit servett / å sen måste du / måste du riva de sönder många gånger (F76)

eh ta en / vit / servett / och // klipp den me händerna / i många delar (S58)

me dina händer du måste bitar /.../ serviett blir bitar (S10)

mh du måste ta en vit serviett / och med dina händer göra många bitar (S54)

va heter de / inte skära inte klippa /.../ de blir små bite (F22)

en servetter vit servetter / eh dra på eh lilla eh bit ne lilla biten / dra ja / helt (P4)

Figur 3. Realisationer av *klippa* och *riva*.

Som kommer att framgå av tabellen nedan använder många informanter det korrekta verbet *klippa* men som framgår av exemplen är informanterna inte alltid medvetna om den exakta betydelsen av verbet trots att det råkar bli rätt: "ta en / inte kniv men den andra". Detta är en onödig precisering. Även om det visar sig att ganska många behärskar *riva*, så tillgrips omskrivningar i flera fall. Ofta förekommer substantiven *del* eller *bit* i dessa, vilket kan tas som ett tecken på att dessa substantiv, som är nära relaterade till fältet diskonnexion, är lättare att lära in än diskonnexionsverben. I tabell 1 ges en statistik över vilka verb som inläraarna använder.

Korrekt verb	Använt verb				
	<i>klippa</i>	<i>skära</i>	<i>riva</i>	Annat	Totalt
<i>klippa</i>	47	3	--	10	60
<i>skära</i>	30	17	--	13	60
<i>riva</i>	7	--	36	17	60

Tabell 1. Inlärnarnas uttryck för diskonnexionsverben

Som framgår använder en stor del av inlärnarna *klippa* när detta verb ska användas. Men många kan inte hålla isär detta verb från *skära*. Exakt hälften använder *klippa* i stället för *skära*. Visserligen finns det några som gör fel i motsatt riktning. Men de är endast tre. *Riva* behärskas av drygt hälften av inlärnarna. Men även i detta fall förekommer det några som använder *klippa* i stället, vilket visar vilken stark ställning detta verb har inom fältet. Personligen blev jag ganska förvånad över resultatet. Även om jag hade förväntat mig att den semantiska kontrasten mellan *skära* och *klippa* skulle neutraliseras, hade jag inte väntat mig att man så entydigt skulle favorisera ett verb. Men resultatet stämmer rätt väl med en allmän tendens, nämligen att inlärnarna favoriserar verb med en tydlig form, då det finns två (eller flera) verb med en nära besläktad betydelse. *Klippa* är ju regelbundet medan *skära* har en mycket oregelbunden böjning. Eftersom det inte finns någon semantisk distinktion i modersmålet på den här punkten, tycks det kännas naturligt för inlärnarna att låta ett verb uttrycka både betydelsen 'skära' och 'klippa'. Som exponent för denna gemensamma betydelse väljer man den ordform som är lättast att uppfatta.

Inom ramen för denna korta uppsats har det endast varit möjligt att kort exemplifiera olika aspekter av tvärspråklig lexikologi men det som sagts räcker förhoppningsvis för att visa hur ett uppmärksammande av tvärspråkliga faktorer kan bidra till en förklaring av hur ordförrådet är organiserat i mänskliga språk.

## Litteratur

- Berlin, B & Kay, P (1969). *Basic color terms*. University of California Press, Berkeley.
- Croft, W (1990). *Typology and universals*. Cambridge University Press, Cambridge.
- Dixon, R M W (1977). Where have all the adjectives gone? *Studies in language* 1: 19-80.
- Greenberg, J (1966). *Language universals with particular reference to feature hierarchies*. Janua Linguarum, Series Minor, 59. Mouton, The Hague.
- Miller, G A & Johnson-Laird, P (1976). *Language and perception*. Cambridge University Press, Cambridge.
- Viberg, Å (1981). *Studier i kontrastiv lexikologi* (ak. avh.). Lättast tillgänglig som: *SSM Report 7-8*. Inst f lingvistik, Stockholms universitet.
- (1984a). The verbs of perception: a typological study. *Linguistics* 21-1.
- (1984b). Fysiska kontaktverb i svenskan. En skiss. I: *Svenskans beskrivning 14*. Lund.
- (1985a). Hel och trasig. En skiss av några verbala semantiska fält i svenskan. I: *Svenskans beskrivning 15*. Göteborgs universitet.



- (1985b). Lexikal andraspråksinläring. I: *SUM-rapport 2*. Centrum för tvåspråkighetsforskning.
- (1986). The study of crosslinguistic lexicology. I: Dahl, Ö (ed.), *Papers from the Ninth Scandinavian Conference of Linguistics*. Inst of Linguistics, Stockholm University.
- (1988). Ordförråd och ordinläring. Några utgångspunkter för en studie av hur ordförrådet lärs in. I: Hyltestam, K & Lindberg, I (red.), *Första symposiet om svenska som andraspråk*. Centrum för tvåspråkighetsforskning.
- (1990). Svenskans lexikala profil. I: *Svenskans beskrivning 17*. Åbo akademi.
- (1991a). Crosslinguistic perspectives on lexical organization and lexical progression. Under utgivning. I: Hyltestam, K & Viberg, Å (eds.), *Progression and regression in language*.
- (1991b). En longitudinell djupstudie av språkutvecklingen. *Utvärdering av skolförberedelsegrupper i Rinkeby. Rapport 4*. Centrum för tvåspråkighetsforskning.

## En databaserad begreppsordbok för den äldre svenskan

Artikeln redovisar ett försök att utarbeta en databaserad, begreppsklassificerad ordbok för den äldre och yngre nysvenskan. Ordboken bygger på Rogets "Thesaurus" (och dess svenska motsvarighet: Bring) och Dahlgrens "Glossarium". Ett av huvudsyftena med ordboken är att den skall vara utgångspunkten i skapandet av en strukturell, synkronisk och historisk ordbok.

### 1. Bakgrund och syfte

I en artikel från 1960-talets mitt skriver Bengt Loman att en uppgift för den framtida historiska lexikografien skulle vara att "utarbeta detaljerade, begreppsligt uppställda ordböcker över en epoks eller en författares ordförråd, varvid synonymiken ägnades en framträdande roll" (Loman 1965:84f). Med datorns hjälp, menar författaren, skulle ett sådant arbete vara möjligt att utföra.

Det projekt, som jag här skall beskriva, syftar till att utarbeta en ordbok av det slag som Loman skisserar, dvs en databaserad begreppsordbok för den äldre och yngre nysvenskan.

En sådan ordbok kommer att vara ett viktigt hjälpmedel för historisk, ordsemantisk forskning, för studiet av betydelsefält, för stilistiska studier rörande ordval och synonymik i äldre texter, komparativa ordstudier (genom sin nära anknytning till existerande begreppsordböcker för engelskan, tyskan, nederländskan, danskan m fl). Även för historiker som arbetar med äldre texter kan den databaserade ordboken vara till hjälp vid texttolkningen.

Projektet har två konkreta utgångspunkter, nämligen C. E. Brings "Svenskt ordförråd ordnat i begreppsklasser", som utkom år 1930 respektive C.F. Dahlgrens "Glossarium öfver föräldrade eller ovanliga ord och talesätt i svenska språket från och med 1500-talets andra årtionde", som utkom i tryck första gången 1914-16.

Projektet anknyter således till de lexikala studier som har sin historiska utgångspunkt i de begreppsordböcker el sk ideologiska ordböcker som tillkom f o m mitten av 1800-talet med P.M. Rogets Thesaurus som förebild: Sanders (1873 ff), Schlessing (1881), Wehrle & Eggers (1967) för tyskan, Harry Andersen (1945) för danskan, Bring (1930) för svenskan m fl) och som så småningom ledde över till elaborerade begreppssystem som exempelvis R. Halligs & W. von Wartburgs Begriffssystem als Grundlage für die Lexikographie (2 uppl 1963):

Till denna typ av ordstudier hör utan tvekan det omfattande forskningsprojekt som tysken Leo Weisgerber står bakom ("Von den Kräften der deutschen Sprache") och som avsatt fyra volymer (1962ff) men också den forskning som Reinhart Koselleck initierat på tysk botten och som går under namnet "Begriffsgeschichte" (Koselleck 1979). Av dessa båda står framförallt den förstnämndes forskning i intim kontakt med den mer strikt språkligt definierade ordsemantiken hos Trier (1934), Öhman (1953) och Oksaar (1958) m fl. På senare

år har denna inriktning också fått avläggare med utgångspunkt i svenskt ordmaterial i Wikströms (1977) och Kukkonens (1989) avhandlingar.

## 2. Material

Mina två utgångspunkter, Bring och Dahlgren, som samtidigt representerar mitt huvudmaterial förtjänar i detta sammanhang en närmare presentation.

### Brings ordbok

Brings ordbok är, vilket också påpekats ovan en "efterbildning" av P. M. Rogets "Thesaurus of English Words and Phrases".

Den inleds med en uppräknig av namnen för de sk "begreppsklasserna" (= 1. Tillvaro, 2. Intighet, 3. verklighet osv), sammanlagt 1000 stycken. Därefter följer i bokens första häft en listning av ord för varje begreppsklass. Orden är där grupperade efter ordklass. I bokens senare häft finns ett alfabetiskt ordnat ordregister som anger till vilken begreppsklass ett visst ord hör. Många ord tillhör flera begreppsklasser.

Brings ordbok saknar helt förklaringar rörande dess syfte. Den ger heller inga läsanvisningar, varken för begreppssystemets eller ordboksartiklarnas uppbyggnad (utöver ordklassindelningen).

Brings uppräknig av begreppsklasserna döljer helt det underliggande sinnrikt utbyggda system för begreppsgrupperingen (och ordgrupperingen) som finns hos Roget.

I själva verket låter sig de 1000 namnen på begreppsklasserna hos Bring inordnas både i en huvudklass och i en under denna ställd sk sektion ('section'). Följande sex huvudklasser finns hos Roget: 1. Abstract relations 2. Space 3. Matter 4. Intellect 5. Volition och 6. Emotion

Hur sektionindelningen (och underavdelningarna till dessa) ser ut framgår av tablån i bl a Kirkpatrick's utgåva av Rogets Thesaurus (1987).

Med en sådan tablås hjälp kan vi alltså sätta in varje begreppsklass hos Bring i ett större sammanhang, få dess klass- och sektionstillhörighet med eventuella underavdelningar.

Brings enkla numrering av begreppsklasserna döljer också den gruppering i kolumner som de fulla systemet i tablån tillhandahåller. I spalterna eller kolumnerna ställs relaterade glosor bredvid varandra.

Detta om själva systemet. Om vi sedan granskar själva ordboksartiklarna som följer varje begreppsklass är orden i dessa som tidigare nämnts, grupperade efter ordklass (s, v, a, se nedan). Varje ordklassparti innehåller en listning av ord. Listningen är inte alfabetisk. Varje listning avslutas med ett semikolon varefter nytt stycke börjar på indragen rad.

Vi tar begreppsklass nr 423 'Ljuskälla' för närmare granskning. En titt i tablån (se Kirkpatrick) anger dess plats i det fulla systemet. Nr 423 'Ljuskälla' hör till huvudklass III 'Matter', sektion III 'Organic Matter' med underavdelning 2 'Sensation' närmare bestämt underavdelning 'Special' närmare bestämt underavdelning 'Light' närmare bestämt underavdelningen 'Light in general' där begreppsklassen står i opposition till 424 'Skugga'. Med andra ord har begreppsklassen med materien att göra (på tyska 'Stoff'), närmare bestämt med den organiska materien (till skillnad från den oorganiska) och den har med allmänt ickespecificerat (dvs ej färger) ljus och synsinne att göra. Ändå mer sammanfattande: vi har med en relativt konkret kategori att göra.

Så här ser begreppsklassen 'Ljuskälla' ut hos Bring:

**423.Ljuskälla;** sol (-skiva), måne, planet, himlakropp, s  
 tjärn-a (-enatt), vintergatan,  
 norrsken, zodiakalljus, komet, stjärnfall,  
 meteor, ljungeld, (korn-, kul-)blix, fosfore-  
 scens, eld (-gnista), flamma, låga, glöd,  
 elektricitet, lyse, gas (-glödljus, -kandelaber,  
 -lampa, -lykta, -lysning, -låga), olj-a (elampa),  
 fotogen (-lampa), acetylen (-gas), (form-,  
 gren-, kalk-, stearin-, talg-, vax-)ljus, dank,  
 fackla, bloss, lunta, pärt-a (-bloss, -klyka,  
 -sticka), marschall;  
 lampa, astral-(bords-)lampa, auer-lampa,  
 (-ljus), lampett, ljus-kastare (-krona, -låga),  
 kron-a (-ljus), lyster, kandelaber (-arm),  
 arm-stake (-ljusstake), ampel, lanterna, (blind-)  
 lykta;  
 fyr (-båk, -torn), blänk-(blink)fyr, båk  
 fyrverkeri, (-pjäs), raket, pyroteknik, ljus-  
 öppning, fönster.  
 v. tända, avbrinna, lysa, fosforescera,  
 öppna.  
 a. (själv-)lysande, elektrisk, stjärnbeströdd,  
 genomskinlig.

Som vi ser har vi 3 stycken eller paragrafer för substantiv, ett stycke för v och ett stycke för a (=adjektiv, adverb mm.). Detta är en ganska liten begreppsklass.

Låt oss sedan titta på det första stycket i s(substantiv)klassen. Det första vi lägger märke till är att den tågordning i vilken orden följer i en paragraf inte är förklarad (vare sig hos Bring eller Roget) ej heller har paragraferna något namn. Tanken bakom ordningsföljden (mellan såväl ord som paragrafer) får varje användare av ordboken själv sluta sig till. Den enda hjälp som finns är styckeindelningen och paragrafens inledningsglosa.

Vi prövar en läsning av den första substantivparagrafen. Det tycks här som om orduppräknningen går från en uppräknning av *kosmiska* ljuskällor : 'sol, solskiva, måne, planet, himlakropp, stjärna, stjärnenatt, vintergatan, norrsken, zodiakalljus, komet, stjärnfall, meteor' över till *mindre, specifika, himmelska fenomen* som 'ljungeld, kornblix, kulblix' och därefter över till *mindre, allmänna* företeelser ej nödvändigtvis av människan åstadkomna ljuskällor som 'fosforescens, eld, eldgnista, flamma, låga, glöd' till av *människor i bruk tagna ljusfenomen* el ämnen 'elektricitet, lyse' (lysämne i allmänhet), 'gas' (varvid företeelser och ting förknippade med gas räknas upp: 'gasglödljus, gaskandelaber, gaslampa, gaslykta, gaslysning, gaslåga'), 'olja' (samma sak här: exempel på tillämpning av olja som ljuskälla: oljelampa) 'fotogen' (med exemplifieringen: fotogenlampa), över till nästa lysämne: 'acetylen' (med exemplet:(acetylengas).

Därefter tycks tanken komma in på *föremål* eller *redskap som människan förfärdigat* för smärre ljuskällor: 'ljus' (av olika slag: 'formljus, grenljus, kalkljus, stearinljus, talgljus, vaxljus, dank'). för att sedan fokusera en *detalj* på ett sådant ljusbringande föremål, det ställe där själva ljuslågan finns: 'vaxstapel, veke'. Tanken glider sedan in på det *ämne* som är den fysikaliska förutsättningen för att frambringa ljus: 'fosfor', och som används vid tillverkning av 'tändsticka'. Vilken roll 'tran' har i detta sammanhang vet jag inte men efter detta ord går

associationen till olika *redskap* som används för att bära eller hålla ljus: 'fackla, bloss, lunta, pärtä' (med tillhörande exempel: 'pärtbloss, pärtklyka, pärtsticka') och 'marschall'.

Den andra substantivparagraf (s2) tycks närmast innehålla en uppräknning av *tillverkade* föremål eller apparater som användes vid belysning.

Den tredje substantivklassen (s3) tycks omfatta något liknande: olika anordningar för att ge ljus, först en sammanhållen grupp 'fyr, fyrbåk, fyrtorn, blänkfyr, blinkfyr och båk' (tornliknande med vägledande uppgifter) därefter 'raket, pyroteknik' (ljus vid speciella tillfällen: högtider) och slutligen 'ljusöppning, fönster' (öppningar i en vägg, tak osv).

Orden i begreppsklassen är naturligtvis samtliga släkt med varandra - på betydelseplanet. Den bästa begreppsliga sammanfattningen på denna släktskap är begreppsklassens "namn". "Ljuskälla" är alltså den minsta betydelsemässiga nämnaren för samtliga ord i samtliga stycken.

Inom varje stycke eller paragraf är orden mer besläktade med varandra än över styckegränserna.

Det inledande ordet i stycket verkar ganska godtyckligt: 'fyr' inleder paragraf 3 och 'fönster' avslutar samma paragraf utan synbar motivering.

Ett stycke tycks beskriva en tankes 'väg'. Det går att se en "mening" en röd tråd i varje stycke även om skälet till att 'vägen' ser ut som den gör, ej står klart. Ett stycke tycks bestå av en eller flera länkar, som relativt enkelt går att sluta sig till och rubricera (på ett sätt som påminner om hur man gör när man ledfamiljsbestämmer innehållet i en text): 1. kosmiska ljuskällor, 2. mindre, specifika, himmelska ljusfenomen 3. mindre, allmänna företeelser ej nödvändigtvis med människan som upphovsman 4. ljusämnen som människan förstått att utnyttja 5. föremål som människan uppfunnit för mindre ljuskällor 6. en viktig detalj på ett sådant föremål 7. ämne som används vid tillverkning av ett redskap som används för att åstadkomma eld ('fosfor') 8. föremål som används för att åstadkomma eld (tändsticka) 9. föremål som brinner och som människan kan bära med sig.

Innehållet i exemplet tycks vidare vara av *encyklopedisk* och ganska *tillfällig* art. I s2 räknas lampor av olika slag upp men ej samtliga t ex finns inte "bilstrålkastare", 'ficklampa' osv med. Exemplet tycks vara viktigare än ett uttömmande svar.

Spännvidden i en paragraf verkar också kunna variera avsevärt. I s1 har vi en stor betydelsemässig spännvidd: innehåller på samma gång 'sol' som 'acetylen' och 'veke'. I s2 har vi en betydligt mindre betydelsemässig spännvidd: ljuskällor tillverkade av människan av en viss typ.

Synonymer (i en mening som exempelvis ordboken Ord för ord använder) tycks återfinnas i ett styckes ordlistning. Ord för ord anger 'bloss, fackla, marschall, pärtä, pärtsticka' som synonymer (likaså 'fyrtorn, fyrhus, fyrbåk').

### Dahlgrens glossarium

Dahlgrens glossarium bygger enligt Evald Ljunggren "på ett ansenligt excerptmaterial, hämtat från bortåt 800 skrifter". Ljunggren anmärker också rörande volymen på ordboken att hela antalet artiklar kan "anslås till över 18000 och språkproven till över 46000" (Dahlgren & Ljunggren 1914-16:IV). Som bekant omspanner Dahlgren tidsperioden från 1500-talet till och med 1700-talet, med tyngdpunkten på litteraturen från femton- och sextonhundratalet, men "lämar åsido allt som överensstämmer med det nutida gängse språkbruket" (ibid.:IV). I sina betydelseförklaringar ger Dahlgren i regel en översättning till en nusvensk glosa. Defintionsmässiga omskrivningar är sällsynta (ibid.:VII). Detta är som det kommer visa sig en klar

fördel för mig i min bearbetning av Dahlgren. Dahlgrens ordförklaringar är enligt Ljunggren "goda och träffande" även om han medger att tolkningen ibland kan bli grov och att det "stundom" händer att "översättningsordet snarare står som en antydning om åt vad håll betydelsen går än ger själva betydelsen" (ibid.:VII).

### 3. Metod

Den första etappen

Utgående från Bring (dvs Rogets klassificering) avser jag att i en första etapp begreppsklassbestämna samtliga ordbetydelser i Dahlgren. Valet av Rogets system har flera fördelar:

- a) orden placeras in i ett "öppet" system som tillåter jämförelser med andra språk (ordförrådet i en rad språk har ordnats efter detta system).
- b) Rogets strukturering är visserligen ingen strikt strukturell betydelsefältanalys men betydelsefälten är tydligt urskiljbara inom begreppsklasserna.
- c) systemet har använts för nysvenskans ordförråd (Bring), något som blir viktigt i den andra etappen i arbetet (se nedan).

Klassificeringen av Dahlgrens ordmaterial skall tillgå på följande sätt:

- 1 Samtliga stycken (efter semikolon och med indragen rad) i Bring skall indexeras (med siffror). Dessa stycken representerar synonymkedjor och begreppsmässigt associativa kedjor av ord underställda ett "språkligt" begrepp (de sk "begreppsklasserna").
- 2 Alla ordbetydelser angivna med "översättningsord" hos Dahlgren skall inordnas i Brings (dvs i stort sett Rogets) system med angivande av siffra för begreppsklass. Undantag är ord i vissa fasta fraser.
- 3 Alla ordbetydelser angivna med översättningsord hos Dahlgren skall förses med indexsiffra för vilket "stycke" de tillhör i Brings styckeuppdelade ordboksartiklar. Samma undantag som ovan. Någon indexering av orden inom ett stycke skall inte göras.
- 4 Det hela skall läggas in i en databas, som även ger de av Dahlgren angivna ordklassetiketterna och etymologierna (utländsk förebild för det lånade ordet samt långivande språk).

Resultatet kommer således bli en begreppsklassificerad och preliminärt betydelsefältsklassificerad (om vi lite vårdslöst skapar till en glosa för att täcka vad ett "stycke" är hos Bring) ordbok i databasform för den äldre och yngre nysvenskan (och ett något senare språkskede, åtminstone en bit in på 1800- talet, dvs Dahlgrens material).

Sammanlagt är det ca 23905 ordbetydelser hos Dahlgren som på detta sätt skall klassificeras.

Det som kommer att saknas i databasen är språkproven och källangivelserna i Dahlgren. Dessa uppgifter får tills vidare hämtas i den "gamla" ordboken.

Källorna i Dahlgren finns oftast angivna i ordbokens "Förteckning på anförda skrifter" med undantag av "en och annan citerad skrift...emedan den till fyllest angifves i citatet". Samtliga i förteckningen namngivna författare skall bestämmas till ålder och födelselandskap. Skrifterna skall också klassificeras vad beträffar ämnesområde eller genre.

Målsättningen här är att fastställa glossariets representativitet vad beträffar beläggkronologin och den ämnes-/genremässiga spridningen.

#### 4. Vad kommer man att kunna använda DBO till efter den första etappen?

Liksom i Brings ordbok skall vi i den databaserade ordboken för äldre svenska kunna söka efter två linjer: från begrepp till ord och från ord till begrepp.

Från begrepp till ord:

Genom att söka utifrån ett begreppsklassnummer kommer vi få samtliga ord tillhörande en viss begreppsklass listade. Vi kan om vi så vill också begränsa sökningen till att gälla enbart en viss ordklass (s, v, a så t ex lista alla substantiv tillhörande begreppsklassen x).

Vi kommer vidare kunna söka ord för vilka kombinationer av begreppsklasser som helst (lista ord för begreppsklasserna x, y, z osv).

Vi kommer också kunna söka efter ord som uppfyller vissa formkriterier inom en viss begreppsklass (lista alla ord inom begreppsklassen x som slutar på -ing).

Från ord till begrepp:

Vi kommer kunna söka begreppsklassen (eller begreppsklasserna) för ett visst ord. Vi kan låta sökningen gälla enbart låneord (lista alla ord som har isländskt ursprung för en viss begreppsklass) eller ord med viss form (vilka begreppsklasser finns antecknade för ord med ett visst prefix, suffix, rotmorfem osv).

Naturligtvis kommer vi kunna använda databasen för helt andra sökningar än de som gäller relationen ord - begreppsklass, exempelvis lista alla ord (utan angivande av begreppsklass) med ett visst suffix, prefix eller rotmorfem (ren morfologisk sökning), lista alla ord som Dahlgren anför som tyska, isländska, latinska osv lånord (etymologisk sökning), lista alla ord som har en viss betydelse, (lista alla ord som hos Dahlgren betyder "anfälla" eller lista alla betydelser för ett visst ord) semantiska sökningar.

En viktig uppgift för DBO blir således att ta fram orden i en viss begreppsklass. I bilaga 1 illustrerar jag med en tänkt sökning i DBO på begreppsklass nr 1 "Tillvaro".

Ett av våra huvudmål kommer att på detta sätt nås: att framställa ett hjälpmedel som förmår ta fram synonymer och betydelsebesläktade ord och sammanställa och strukturera dessa i begreppsklasser och paragrafer.

Notera dock att de två spalterna inte representerar två olika tidsepoker, som är klart tidsmässigt avgränsade från varandra. Många av vänsterspaltens ord har ju funnits även under äldre nysvensk tid och högerspaltens ord representerar inte den äldre nysvenskans samtliga ord utan enbart de ord (och ordbetydelser) som försvunnit ur språket.

## 5. Projektets andra etapp: att nå fram till den "fullständiga" strukturella, synkrona, historiska ordboken

En andra etapp i projektarbetet är att nå fram till den "fullständiga" strukturella, synkrona, historiska ordbok (som Loman talar om) för äldre svenska, som förmår dra tidgränser i den historiska utvecklingen, som t ex klart kan separera den äldre nysvenskan från nusvenskan eller som förmår avgränsa 1500-talsspråket från 1600-talsspråket osv.

För att skapa en fullständig, strukturell, historisk ordbok för den äldre och yngre nysvenskan som förmår göra synkrona avgränsningar måste vi *tidsbestäm* samtliga Brings glosor och i ordboken *inkorporera* respektive *utesluta* de ord (och ordbetydelser) som introducerats *före* respektive *efter* periodgränsen mellan yngre nysvenska och nusvenska (ca år 1800). Tidsbestämningen skall ske med utgångspunkt i SAOB.

I bilaga 2 skisserar jag hur en sådan tidsbestämning skall gå till. Jag ger som exempel tidsbestämningen av begreppsklassen 'Tillvaro' (av utrymmesskäl inskränkt till att gälla endast substantiven). Några ord måste dock utgå av det enkla skälet att de ännu ej omfattas av SAOB.

I detta sammanhang kan vi återigen aktualisera det som Loman påpekade i sin artikel om utarbetandet av en strukturell, historisk ordbok som förmår göra synkrona avgränsningar, nämligen att ett sådant arbete måste gå hand i hand med utvecklade och förfinade semantiska metoder. För Loman innebär detta huvudsakligen i att ordens semantiska mönster fastställs genom studiet av kontexterna utifrån från två "formella kriterier" - substitution och distribution.

Detta steg tar vi emellertid inte i det projekt som här presenteras utan nöjer oss med att hävda att grunden för ett sådant arbete i och med existensen av den databaserade begreppsordboken härmed är lagd.

## 6. Sammanfattning

Vi kommer att som resultat av den första etappens arbete få en databaserad begreppsordbok för den äldre och yngre nysvenskan. Denna version av ordboken är dock "ofullständig" eftersom den endast omfattar Dahlgrens ordmaterial.

Den andra etappens arbete innebär att fastställa vilka nusvenska glosor (i Bring) som existerat i en viss betydelse i den äldre och yngre nysvenskan. Denna version av ordboken kan betraktas som en "fullständig", strukturell, historisk ordbok, som förmår lägga grunden till en synkronisk betydelseindelning av orden (exv 1500-talets, 1600-talets, 1700-talets synonymik).

En sådan ordbok kommer bli att göra det möjligt att med Lomans ord "blottlägga en semantisk struktur i ett gånget skede" och "se hur ett visst begrepp eller begreppskomplex varit semantiskt indelat genom ord, och att se vilka olika betydelsevarianter detta eller dessa ord haft" (Loman 1965:85).

Målet att jämföra "semantiska strukturer från olika skeden i ett språks historiska utveckling" (Loman 1965:86) har i och tillkomsten av den databaserade begreppsordboken därmed nåtts.



## Litteratur

- Andersen, H. 1945. *Dansk Begrebsordbog*. Köbenhavn.
- Bring, S.C. 1930. *Svenskt ordförråd ordnat i begreppsklasser*. Stockholm.
- Dahlgren, F.A. 1914-16. *Glossarium öfver föråldrade eller ovanliga ord och talesätt från och med 1500-talets andra årtionde*. Lund.
- Hallig, R. & W. von Wartburg. 1963. *Begriffssystem als Grundlage für die Lexigraphie*. 2 uppl.
- Koselleck, R. (ed). 1979. *Historische Semantik und Begriffsgeschichte*. Stuttgart.
- Kukkonen, P. 1989. *Från konst till vetenskap*. Helsingfors.
- Loman, B. 1965. Nya uppgifter för historisk lexikografi. I: Ekbo, S & B. Loman, Vägledning till Svenska Akademiens ordbok. *Skrifter utg av Nämnden för svensk språkvård* 32. Stockholm.
- Oksaar, E. 1958. *Semantische Studien im Sinnbereich der Schnelligkeit*. Stockholm.
- Roget's *Thesaurus*. 1987. new edition by B. Kirkpatrick. London & Bath.
- Sanders, D. 1873ff. *Deutscher Sprachschatz, geordnet nach Begriffen*.
- Schlessing, A. 1881. *Der passende Ausdruck*.
- Thesaurus of English words and phrases* by P.M. Roget. 1953. Enlarged by J.L. Roget. New edition revised and enlarged by S.R. Roget. London m. fl.
- Trier, J. 1934. Das sprachliche Feld. *Neue Jahrbücher für Wissenschaft und Jugendbild* 10, 428-449.
- Weisgerber, L. 1962ff. *Von den Kräften der deutschen Sprache*, I, II, III, IV. Düsseldorf.
- Wehrle, H & H. Eggers. 1967. *Deutscher Wortschatz*. 13 uppl.
- Wikström, K. 1977. *Begreppsfältet kroppsövning*. Semantiska studier i svenskt idrottsspråk. Acta Universitatis Tamperensis Ser A Vol 83. Tammerfors.
- Öhman, S. 1953. Theories of linguistic Field. *Word* 9, 123-134.

## Bilaga 1

Substantivparagraferna i begreppsklassen "Tillvaro" i en tänkt sökning i DBO

'Tillvaro' hör till huvudklassen I 'Abstract relations'. Denna huvudklass är, som framgår av tabblån, indelad åtta sektioner (Existence, Relation, Quantity, Relation, Order, number, Change och Causation). 'Tillvaro' hör till sektion 1 'Existence', som i sin tur är indelad i 4 grupper (Ab-stract, Concrete, Formal, och Modal). 'Tillvaro' hör till underavdelningen 'Abstract' av 'Existence'." Begreppsklassens relaterade begreppsklass är begreppsklass nr 2. 'Intighet'.

Begreppsklassen 'Tillvaro' har 5 "stycken" eller paragrafer för s, 3 "stycken" för v och 5 "stycken" för "a". Här återges endast s-paragraferna. I spalten för Dahlgren har orden listats alfabetiskt.

## Bring

## Dahlgren

s1 Tillvaro; existens, vara, varande varelse, tillvarelse(-form), befintlighet, före-fintlighet (-komst), substantialitet, realitet, reale, verklighet, aktualitet (-isation, -isationsprocess, faktiskhet, (orubbligt) faktum, fakta, objektivitet;	sanning (=verklig- het); vare (=varande); varelse (till- varelse); varelse (=varande);
s2 värld, världsalltet, skapelse, ontologi, (tingens) väsen, väsentlighet, företeelse, fenomen, sätt att te sig;	kreatur (= skapelse skepsel/skäpsel (= skapelse) underhålle (=uppe- hålle) varelse (=väsende)
s3 (själv-)bestånd, substisens, fort- varo(-komst), liv, själv-uppehållelse (-bestämning), uppehålle, utkomst;	beständelighet (= bestånd), fordenskap (=uppehålle); hals (=liv) lefnad (=li) lefvande (=liv) lifaktighet (=liv); lifnad (=liv); själ (=liv); ståndighet (bestånd) underhålle (uppehålle);
s4 när-(sam-)varo, koexistens, sam- verkan;	närvist (=närvaro);
s5 riktighet, (naken) sanning, allvar (och icke skämt).	allvarhet (=allvar); allvarlighet (=allvar); sann (=sanning);

**Bilaga 2**

Tidsbestämning av substantiven i paragraf s1 i begreppsklassen "Tillvaro".

(D= Dahlgren, a3, r5 osv återger SAOB:s indexering av ordets betydelser; I vissa fall saknas exakt årtal men kan ersättas av uppgift om skrift (exv Gustav I:s registratur (GIReg), Gustav II Adolfs skrifter (GIIA)). Så långt det varit möjligt har uppgifter hämtats från SAOB.

**S1**

sanning (=verklighet)	GIreg, D
vare (=varande)	GIreg, D
varelse (=tillvarelse)	1541, D
-----	
aktualitet a3	1652, SAOB
realitet r5	1670, SAOB
realitet r3	1680, SAOB
realitet r6	1689, SAOB
ej exakt årtal för 'varelse' (=varande; A. Oxenstierna)	
-----	
existens, e1	1718, SAOB
reale, r3	1749, SAOB
faktum f2	1755, SAOB
objektivitet	1795, SAOB
realitet r1	1799, SAOB
-----	
befintlighet	1807, SAOB
faktiskhet	1823, SAOB
aktualitet a1	1847, SAOB
förekomst	1858, SAOB
aktualisation	1861, SAOB
realitet r4	1868, SAOB
realitet r2	1880, SAOB
förefintlighet	1885, SAOB
reale r5	1894, SAOB
aktualitet a2	1895, SAOB

## Varians og distinktivitet

Indlæggets tema er lemmatiseringen af homonyme leksemer i historisk leksikografi, specielt ældre nyhøjtysk (tysk 1350-1650). Det undersøges, om den homonyme konstellationstype *heterografi*, *men homofoni* findes belagt i ældre nyhøjtysk, samt hvorledes denne type behandles ved lemmatiseringen i de eksisterende ældre nyhøjtyske systemordbøger (Götze og Anderson/Goebel/Reichmann). Eftersom det viser sig, at der ikke i nogen af disse ordbøger tages adækvat hensyn til den, drøftes det, om den alligevel ud fra teoretiske overvejelser af sprogsociologisk art burde repræsenteres, og der anvises forslag til, hvorledes der kunne lemmatiseres inden for rammerne af "det bedste af alle eksisterende ældre nyhøjtyske grafemsystemer".

### Lemmatiseringsproblemer i historisk tysk leksikografi

Det er velkendt, at de tyske grammatikere i det 17. og 18. århundrede i forbindelse med fastlæggelsen af den nyhøjtyske skriftsprogsnorm bidrager til den sproglige monosemeringsproces ved bl.a. at kræve, at der på udtryksplanet indføres grafiske differentieringer for (partielt eller totalt) homonyme leksemer. Således anfører Schottelius i 1641 en længere række eksempler på enslydende ord med forskellig grafisk realisation<sup>1</sup>, og Gottsched opstiller i sin *Deutsche Sprachkunst* i midten af det 18. århundrede følgende ortografiske regel: "Wörter verschiedener Bedeutung, und die nicht voneinander abstammen, unterscheidet man, so viel möglich ist, durch die Buchstaben".<sup>2</sup> Hermed indledes samtidig traditionen for de såkaldte homonym-ordbøger, der - for så vidt angår ordbøgernes antal - når sit højdepunkt i det 19. århundrede, men som i øvrigt stadig varer ved.<sup>3</sup> Homonymordbøgernes formål består typisk i at yde brugerne en håndsregning ved retstavningen, og et eksempel er indeholdt i titlen på Wolfgang Mentzups homonym-ordbog fra 1971: "malen oder mahlen? Gleichklingende, aber verschieden geschriebene Wörter".

### Homonymdifferentiering i nyere tysk ortografi

Trods Gottscheds ovenfor citerede krav er den grafiske homonymdifferentiering dog aldrig blevet systematisk gennemført i tysk<sup>4</sup> (man kan blot tænke på ordbøgernes homonymindekser ved leksemer som *Bremse*<sup>1</sup>, *Bremse*<sup>2</sup> og *Schloß*<sup>1</sup>, *Schloß*<sup>2</sup>). En optælling af heterografe homofoner i Duden Rechtschreibung viser således, at fænomenet i moderne tysk begrænser sig til 24 leksem-par af typen *Seite/Saite*, *Lerche/Lärche*, *Moor/Mohr*, *Leib/Laib* og *Weise/Waise*.

Mange af differentieringsforslagene hos Schottelius og Gottsched har vundet indpas i moderne tysk ortografi, og det siger sig selv, at de ikke volder leksikografen særlige vanskeligheder i ordbøger over nutidssproget, idet f.eks. *Saite* og *Seite* behandles på to

forskellige steder i ordbogen. Det uproblematiske skyldes som bekendt, at den tyske norm siden den anden ortografiske konference i Berlin 1901 kodificeres i Duden Rechtschreibung, hvor graden af grafisk varians for øvrigt er begrænset til dobbeltskrivninger af ganske få fremmedord (f.eks. *Corpus* vs. *Korpus*), hvilket nemt klares ved et system af henvisnings- og hovedlemmaer ansat efter leksikografens skøn.

### Varians og distinktivitet i ældre tysk: Et (meta-)leksikografisk problem

Så enkelt er det dog langtfra, når det gælder lemmatiseringen i de synkront-semasiologiske ordbøger inden for den historiske leksikografi (hvis objekt ifølge nyere metaleksikografisk forskning defineres som sprogrtrin før ca. 1850)<sup>5</sup>. Disse historiske sprogrtrin er nemlig karakteriseret ved en høj grad af ikke mindst en grafisk varians, som er betinget først og fremmest af manglen på ledevarietet, men som igen skyldes politisk-kulturelle forhold (Tysklands mangel på samling og kulturelt centrum) og i nogen grad også datidens æstetiske forestillinger (i én og samme tekst blev ét ord hyppigt skrevet forskelligt, dersom det forekom flere gange) samt endog af økonomiske spekulationer (skriverne fik - en tid - betaling pr. linje, hvilket foranledigede overflødige konsonantophobninger af typen *vwndt* og *vuherdhenn* for *und* og *werden*).

At anføre samtlige varianter ville ikke mindst på grund af ordbogens omfang og overskuelighed selvsagt være utænkeligt. Heraf følger, at den historiske leksikografi må udvikle en metode, hvorved de mange varianter kan overvindes til fordel for repræsentative og samtidig brugervenlige lemmer<sup>6</sup>. I det følgende vil jeg koncentrere mig om den del af den historiske leksikografi, der har det ældre nyhøjtyske system (dvs. højtysk 1350-1650) som objekt, og jeg vil nøjes med at behandle spørgsmålet om den grafiske homonymdifferentiering i lyset af de forhåndenværende ordbøgers lemmakonceptioner: Forekommer typen heterografi/homofoni også i ældre nyhøjtysk, eller findes den først belagt fra Schottelius? Hvis den forekommer: Hvorledes bliver der taget hensyn til den ved lemmatiseringen i de eksisterende ordbøger?

Med andre ord: Hvornår er to grafier *varianter* og hvornår er de *distinktive enheder*?

### Lemmatiseringsmodeller for ældre nyhøjtysk

Der findes kun to ordbøger, som har hele det ældre nyhøjtyske system som objekt, nemlig dels Alfred Götze: *Frühneuhochdeutsches Glossar* (1912 med senere optryk), dels Anderson/Goebel/Reichmann: *Frühneuhochdeutsches Wörterbuch*, et metaleksikografisk meget velfunderet og internationalt anerkendt værk.<sup>7</sup>

Som ovenfor antydnet er lemmatiseringen i historisk leksikografi en abstrakt proces, og det skal fremhæves, at når der tales om homonymi og forskellige homonymi-konstellationstyper, så sker dette altid *relativt* til den abstrakte lemmatiseringskonception, der så at sige træder i stedet for den ikke-eksisterende fællesnorm. Derved får lemma-konceptionerne en vis form for plasticitet, som muliggør, at der kan argumenteres for nye leksikografiske handlingsskridt inden for rammerne af allerede eksisterende modeller (hvilket skal forsøges i det følgende). Først må der altså ses nærmere på de nævnte to ordbøgers lemma-konceptioner, dernæst på, hvorledes homonymdifferentieringen eventuelt måtte blive behandlet der.

### Lemmatiseringen hos Götze

Der findes ingen offentliggørelse af de principper, lemmatiseringen hos Götze hviler på. Man må derfor ty til de få bemærkninger i ordbogens indledning, når man vil prøve at finde ud af, hvilke tanker der ligger bag. F.eks. fremgår følgende af Götzes indledning<sup>8</sup>:

- 1 den nyhøjtyske diftongering af middelhøjtysk *ī, ū* og *iu* [y:] er gennemført, således at der konsekvent lemmatiseres *ei, au* og *eu*;
- 2 ved den nht. monoftongering af de gamle diftonger mht. *ie, uo, ūe* til *ī, ū, ū* er lemmatiseringen - uvist af hvilke årsager - ikke konsekvent;
- 3 der gives afkald på grafien *ai* med den begrundelse, at der ellers ville opstå skriftbilleder, som havde stået i et modsætningsforhold til både den middelhøjtyske og nyhøjtyske norm<sup>9</sup>.

Ser man nu på de enkelte lemmer hos Götze, bliver billedet endog mere uklart, end man kunne forvente ud fra punkt 2 ovenfor. Ved lemmaet *wīß* 'Strafe' er den nht. diftongering trods oplysningerne i indledningen *ikke* gennemført, hvilket dog modsat er tilfældet ved lemmaet *weise* f. 'Handlungsweise, Verfahren'.

Et potentielt tilfælde af homonymi mellem to leksemer, i hvilke den betonede vokal begge steder går tilbage på den middelhøjtyske lange monoftong *ī* undgås altså ved, at grafiske realisationer fra diasystematisk betragtet forskelligt forankrede tekster på ukritisk (og også helt uacceptabel) vis paralleliseres. I sin indledning anførte Götze, at han giver afkald på grafien *ai*. Man skulle derfor forvente, at Götze nu ville anføre to lemmer med homonymindekser: *weise<sup>1</sup>* og *weise<sup>2</sup>* for hhv. "måde" og "forældreløst barn". Dette er imidlertid ikke tilfældet, idet Götze - som er meget selektiv - slet ikke anfører et lemma *weise* i betydningen "forældreløst barn", men vælger i stedet kun at anføre den øvretyske diminutiv *weisel*, hvorved altså endnu et tilfælde af homonymi undgås.

Faktisk lykkes det Götze takket være en yderst inkonsistent (og dermed som nævnt videnskabeligt uacceptabel) lemmatisering kombineret med en høj grad af selektivitet at suggerere et næsten homonymifrit system for ældre nyhøjtysk. Spørgsmålet om, hvorledes den grafiske homonymidifferentiering behandles hos Götze, kan ikke umiddelbart besvares, thi som antydnet mangler der en tertium comparationis, dvs. den konstante basis, i relation til hvilken homonymidifferentieringen eventuelt kunne være forekommet. Heroverfor kunne man nu hævde, at lemmatiseringen *weise* ("måde") over for *wīß* ("straf") er et eksempel på homonymdeling, men da disse leksemer jo ikke tilhører samme dialekt, er de slet ikke kommensurable.

Det må konkluderes, at lemmatiseringen hos Götze ikke sker efter en gennemreflekteret, konsistent model, men snarere synes at bestå i et rodsammen af forskellige dialekter, og inden for rammerne heraf er det slet ikke muligt at tage stilling til konstellationen heterografi/homofoni, som set i forhold til Götzes mange helt elementære problemer bliver et luksus-spørgsmål.

### Lemmatiseringen hos Anderson/Goebel/Reichmann

Er lemmatiseringen hos Götze vilkårlig og ureflekteret, gælder det modsatte for den anden af de ældre nyhøjtyske systemordbøger, nemlig Anderson/Goebel/Reichmann: Frühneuhoch-

deutsches Wörterbuch, til brug for hvis lemma-konstruktion der foreligger et såkaldt *ideelt grafemsystem*<sup>10</sup>. Der går her ud fra, at ældre ældrenyhøjtysk (dvs. ældre nyhøjtysk før ca. 1500) er karakteriseret ved totalt fravær af en ledevarietet, dvs. ved en horisontal sideordning af varieteter, som i yngre ældrenyhøjtysk (fra ca. 1500) udsættes for en stadig tiltagende vertikalisering ved det nyhøjtyske standardsprog.

Konstruktlemmaet defineres (f.eks. modsat i moderne dialektordbøger, hvortil der kan drages paralleller - men *kun* til situationen i ældre ældre nyhøjtysk!) sprogsociologisk som en mereværdig form, således som den ideelt forekommer i *østmellemtysk/østøvretyisk i yngre ældre nyhøjtysk* - en form, som i parentes bemærket med hensyn til fonologien altid stemmer overens med det fonologiske system i nyhøjtysk. De frie varianter skaffes af vejen, idet der først etableres en ideel fonemfølge og dernæst en tilsvarende ideel grafemfølge (= lemmaet).

Til fonemet ænht. /ei/ i den fonologiske idealmodel, som ligger til grund for den grafematiske, svarer nu en række varianter, f.eks. *ei, ey, eÿ, ai, aÿ, eih, aih* osv., og eftersom den bairisk-svabiske *ei:ai*-opposition (for mht.  $\hat{i}$  *ei* og for mht. *ei ai*) udelades som sociologisk mindreværdig, sammenfattes varianterne til grafen *ei*. Da lemmatiseringsmodellen opfatter grafemerne som relative til fonemerne, kan der ikke umiddelbart inden for rammerne af modellen tages hensyn til enheder, som er semantisk distinktive på det grafematiske niveau, uden at der er en tilsvarende opposition på det fonematiske niveau (altså typen *1 fonem, 2 grafemer*), og i indledningen til Frühneuhochdeutsches Wörterbuch hævdes det da også, at den grafiske homonymflugt ligger senere end ældre nyhøjtysk, dvs. at den først forekommer fra Schottelius og frem<sup>11</sup>.

Det skal nu diskuteres, om denne påstand er rigtig, og - hvis den ikke er det - under hvilke omstændigheder det beskrevne idealgrafemsystem, som jeg vælger at lægge til grund som *det bedste af alle eksisterende ældre nyhøjtyske idealgrafemsystemer*, kunne omfatte den pågældende homonymitype.

### Var der en grafisk homonymdeling i ældre nyhøjtysk?

Som nævnt indledningsvis er den grafiske homonymflugt aldrig blevet systematisk gennemført i skriften i tysk - en flugt foreligger principielt, blot der er én der flygter, så kan vi finde blot ét homonympar af typen heterografi/homofoni, eksisterer den, og der må i givet fald tages leksikografisk stilling til den. I det følgende vil jeg opstille den hypotese, at den ideelle fonemfølge ænht. *v-ei-z-a/* er potentielt homonym og dermed også potentielt homoe eller heterograf. Der skal ses på, hvorledes denne fonemfølge er realiseret i forskellige overleverede kilder og på, om der kan findes eksempler på den grafiske homonymdeling.

I Meister Fabritius' *Büchlein gleichstimmender wörter, aber vngleichs verstandes* (1532) finder vi en metasproglig begrundelse for fænomenet (der dog - hvilket ikke skal kommenteres yderligere her<sup>12</sup> - på ukontrollabel vis sammenblandes med fænomenet paronymi, dvs. den omvendte konstellation homografi/heterofoni af typen *ra:ten* vs. *ratten*). Meister Fabritius anfører dog, at der er en grafisk opposition mellem leksemerne *weiss, weyss* og *waiss*, men desværre anfører han ikke disse formers betydning. I *Schryfftspegell* (Köln 1527) finder man dog under overskriften *Gelich stymende worther vnd vngelicher bedeutung, wie man die und der gleich vnderscheydlich vnd verstendlich schriben sal, verba equisonantia genant* den højst bemærkelsesværdige eksempelsætning *Cuntz weyss hat eyn andre wyesse wyesse flissen tzu machen*, som er yderst interessant, eftersom både *wyesse* og *weysse* går tilbage på mht. former med den lange monoftong  $\hat{i}$ , og fordi *Schryfftspegell* altså her i stedet

for enten at beholde monoftongen i begge ord eller diftongere dem begge, da altså funktionaliserer træk dels fra neder-, dels fra højtysk med henblik på homografdeling (en fremgangsmåde, som vist for øvrigt også kendes fra moderne norsk rigsmål, hvor der skelnes mellem *ben* og *bein*, *sten* og *stein*?). Denne opposition kan dog, hvor interessant den end er, og selv om den er metasprogligt begrundet i Schryfftspegell, ikke umiddelbart overføres på den ideelle lemmatiseringsmodel, eftersom fænomenet begrænser sig til den helt specielle sproglige situation i Köln efter det ribuariskes tilbagegang i det 16. århundrede og sammenstødet af flere forskellige andre dialekter, og eftersom den ikke er en sprogsociologisk mereværdig form, således som denne blev defineret ovenfor.

En excepcion, jeg har foretaget af ca. 50 ældre nyhøjtyske kilder sammenstillet alsidigt ud fra diakrone, diatopiske, diastratiske og diasituative kriterier, lader følgende varianter forekomme:

for nht. *weise* (dansk 'klog'):

*weisen, Wysin, wysin, weizen, weÿssenen, weise, weiser, wiser, wîs, wîser, weyse, Weise, weyße*

for nht. *Weise* (dansk 'måde'):

*wyß, wise, weis, wyse, weyse, weise, weyß, Weyse*

for nht. *Waise* (dansk 'forældreløst barn'):

*weise, waisen, weisen, Weysen, Wâise, waisse, waysen, wâysen, Waisen*

for nht. *verwaist* (dansk 'forældreløs'):

*waise, waisse*

Nogle eksempler på, at der i kilderne ofte er grafisk forskel på disse leksemer, uden at der derfor foreligger eksempler på konstellationen heterografi/homofoni:

sie solten schaurm und schirmen witwen und *waisen*; gancz in geistlicher *weise*; doctores, *weise* prister (Die Gedichte des Michel von Beheim, nordøvretyisk, 2. halvdel af det 15. århundrede); heterografien skyldes her en fonetisk forskel i bayersk og svabisk på *ei* og *ai* (ei for mht. î; ai for mht. ei). Altså: ikke homofoni!

megede vnde wetewen vnde *weisen*; jn der silben *wise* nimt ouch ein ieslich man; sine *wise* rede (Schwabenspiegel, nordtysk/mellemtysk, ca. 1410). Da den middelhøjtyske fonologi er bibeholdt (î og ei er to fonemer), foreligger der ikke homofoni.

Luthers bibeloversættelse er interessant for vort formål, idet der i Berliner Handschrift 1523/24 forekommer homofone homografer, der i de senere Wittenberg-tryk fra 1531 og 1545 differentieres til heterografer:

Berliner Handschrift 1523/24: du bist der *weysen* helffer; dem narren gefellet seyne *weyse* wol; das zeugnis des herrn ist gewis vnd macht *weyse* die albern (citatene stammer fra salmernes og ordsprogenes bøger)



Wittenberg-tryk 1531 hhv. 1545: du bist der *waisen* helffer; dem Narren gefelt seine *weise* wol; das zeugnis des Herrn ist gewis, vnd macht die Albern *weise*.

At der her ikke er tale om en dialektal særhed, men om konstellationen heterografi/homofoni, ses af den oprindelig ens skrivemåde i håndskrifterne.

### Homonymdifferentiering og lemma-konstruktion: Mulige leksikografiske handlingsskridt

Eksemplet hos Luther *weise:weise* er altså et bevis for homografdelingens eksistens i ældre nyhøjtysk. Overflødig at bemærke er, at Luther må karakteriseres som en "sprogsociologisk mereværdig skriver" forankret i østmellemtysk yngre ældre nyhøjtysk, hvorfor det sprogsociologiske kriterium for, at der skal tages hensyn til formen ved lemmatiseringen, er opfyldt. Spørgsmålet er dog nu, hvorledes man derefter konkret kan tage hensyn til den, eftersom det ideelle grafemsystem jo ikke accepterer autonomt-grafematiske differentieringer uden pendant på fonem-niveau (fonologisk er der jo ikke forskel på *weise* og *weise*).

Man kunne løse problemet på i hvert fald to måder. Under alle omstændigheder må der afviges fra lemmatiseringsforskriften /ei/ > [ei], nemlig på grundlag af det ovenfor skitserede sprogsociologiske kriterium. Dette kunne gøres ved at ansætte to hoved-lemmaer, nemlig dels *weise* ('måde' etc.), dels *weise* ('forældreløst barn' etc.), men mindre radikalt kunne man afvige fra den relationale grafem-model (som ordbogsbrugeren måske netop har lært at gennemskue) ved at ansætte *weise* som hovedlemma i begge tilfælde og ved betydningen 'forældreløst barn' *weise* som sidelemma (og henvisningslemma!), dvs. en dobbeltlemmatisering *weise, weise* med en henvisningsstruktur *weise s. weise, weise*.

### Homonymdifferentieringen i diakront perspektiv

Det er under alle omstændigheder vigtigt at fastslå, at den grafiske homonymdeling forekommer også i ældre nyhøjtysk, og at frie *varianter* undertiden (som hos Luther) funktionaliseres semantisk og altså i visse tilfælde bliver *distinktive enheder*.

Ordbogen bør afspejle dette fænomen, thi det lader os erkende en vigtig del af monosemeringsprocessen og er tillige interessant i kulturhistorisk sammenhæng, idet Luther her funktionaliserer træk fra en anden dialekt end sin egen (her bairisk/svabisk), vel sagtens ud fra den tankegang, at også dialekt-talende skal føle sig repræsenteret i ledevarieteteten.

Omfanget af homografi-differentieringen i ældre nyhøjtysk har vi ikke kunnet gå ind på. Når der tages hensyn til fænomenet i forbindelse med lemmakonstruktionen, vil det dog til sidst fremgå af den færdige 'Frühneuhochdeutsches Wörterbuch'.

Jeg håber, at det er lykkedes mig ikke blot at give et indtryk af et interessant problem af generel betydning vel både for den tyske leksikografi og for historisk leksikografi i det hele taget, men også at anskueliggøre den vanskelige lemmatiseringsproces ved et eksempel på den grå zone mellem varians og distinktivitet, og at lemmatiseringsmodellens plasticitet muliggør en vis grad af argumentatorisk kreativitet hos de enkelte medarbejdere på et stort ordbogs-projekt.

### Noter

1. F.eks. skelner Schottelius mellem *weise* ('modus') og *weise* ('orphanus'). Dette ordpar er hos Bußmann (1990, 2. udg., s. 313) anført som parade-eksempel på grafisk homonymiflugt.
2. Sprachkunst. I. Theil, II. Hauptstück, V. Regel.
3. En oversigt over tyske homonymordbøger findes hos Kühn (1978, s. 45-46). Disse ordbøger, hvis formål som oftest er et pædagogisk, indeholder flest eksempler på partielle homonymer (fx [ich] *weiß* vs. [die Farbe] *weiß*, hvis grundformer er forskellige (*wissen* vs. *weiß*)).
4. Jfr. f.eks. Nerius (1989, s. 460).
5. Jfr. Reichmann (1984, s. 460).
6. Jfr. Wiegand (1984, s. 569).
7. Ordbogen skal efter planen komme i 12 bind. Indtil nu er bd. I (*a-äpfelkern*) udkommet. Bd. 7 (*i-c/k*) udarbejdes i København (ansvarshavende redaktør: V. Winge). De øvrige bind udarbejdes i Heidelberg, Münster, Newcastle og Lubbock (Texas, USA).
8. Götze (1967), s. IX-X.
9. Dette argument passer netop ikke for homgraf-delte ord (f.eks. *Saite/Seite*) samt ordet *Kaiser*.
10. Det ideelle grafemsystem er indgående beskrevet i en artikel af Anderson/Goebel/Reichmann (1981).
11. Reichmann (1986, s. 74).
12. Jeg henviser til min artikel Mogensen (1992), hvor der er anført en længere række eksempler.

### Litteratur

- Anderson, Robert R./Ulrich Goebel/Oskar Reichmann. 1981. Ein idealisiertes Graphemsystem des Frühneuhochdeutschen als Grundlage für die Lemmatisierung frühneuhochdeutscher Wörter. I: *Studien zur neuhochdeutschen Lexikographie I (=Germanistische Linguistik 3-4)*: 53-122. Hildesheim.
- Anderson, Robert R./Ulrich Goebel/Oskar Reichmann (udg.). 1989. *Frühneuhochdeutsches Wörterbuch*. Bd. 1: *a-äpfelkern*. Berlin/New York.
- Besch, Werner/Oskar Reichmann/Stefan Sonderegger (udg.). 1984. *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. 2 Halbbände. Berlin/New York. (Handbücher zur Sprach- und Kommunikationswissenschaft 2,1; 2,2).
- Bußmann, Hadumod. 1990. 2. udg. *Lexikon der Sprachwissenschaft*. Stuttgart. (Kröners Taschenausgabe; Bd. 452).
- Eyn nutzlich Buchlein etlicher gleich stymender worter Aber vngleichs verstandes / denn angenden deutschen schreybschülern / zu gut mit geteylt /Durch Meister Hanssen fabritiū Rechenmeister vnd deutscher schreyber zu Erffurth*. 1532. Erfurt.

- Gille, Hans/Ingeborg Spriewald (udg.). 1968-77. *Die Gedichte des Michel von Beheim. Nach der Heidelberger Hs. cpg 334 unter Heranziehung der Heidelberger Hs. cpq 312 und der Münchener Hs. cqm 291 sowie sämtlicher Teilhandschriften*. Berlin. (Deutsche Texte des Mittelalters 60; 64; 65,1; 65,2).
- Gottsched, Johann Christoph. 1762. 5. udg. *Vollständige und Neuerläuterte Deutsche Sprachkunst, Nach den Mustern der besten Schriftsteller des vorigen und itzigen Jahrhunderts abgefasst* (...). Leipzig.
- Götze, Alfred. 1967. 7. udg. *Frühneuhochdeutsches Glossar*. Berlin. Erste Auflage 1912.
- Grosse, Rudolf (udg.). 1964. *Schwabenspiegel. Kurzform. Mitteldeutsch-niederdeutsche Handschriften*. Weimar. (Monumenta Germaniae Historica. Fontes Iuris Germanici Antiqui, Nova Series; 5).
- Kühn, Peter. 1978. *Deutsche Wörterbücher. Eine systematische Bibliographie*. Tübingen. (Reihe Germanistische Linguistik RGL; Bd. 15).
- Luther, Martin: D. Martin Luthers Werke. Kritische Gesamtausgabe. *Die deutsche Bibel*. 12 Bände. Weimar 1906-1961.
- Mentrup, Wolfgang. 1971. *mahlen oder malen? Gleichklingende, aber verschieden geschriebene Wörter. In Gruppen dargestellt und ausführlich erläutert*. Mannheim. (Duden Taschenbücher, Bd. 13. Sonderreihe zum Großen Duden).
- Mogensen, Jens Erik. 1992. Heterographie und Homophonie in Frühneuhochdeutschen. Ein Beitrag zur gesamtsystembezogenen Lemmatisierung in der Lexikographie. I: *Zeitschrift für germanistische Linguistik (ZGL)*, Heft 20:64-81. Berlin/New York.
- Nerius, Dieter. 1989. 2. udg. *Deutsche Orthographie*. Leipzig.
- Reichmann, Oskar. 1984. Historische Lexikographie. I: Besch/Reichmann/Sonderegger (udg.): 461-492.
- Reichmann, Oskar. 1989. Lexikographische Einleitung. I: Anderson/Goebel/Reichmann (udg.): 10-164.
- Schottelius, Justus Georg. 1663. *Ausführliche Arbeit von der Teutschen Haupt Sprache*. Braunschweig.
- Schryfftspegell. 1527. *Formulare vñ duytsche Rhetorica, ader der Schryfftspiegel ghenant des neuen stylums vñ pracktiken tzo Missiuen etc.* (...). I: Müller, Johannes: *Quellenschriften und Geschichte des deutschsprachlichen Unterrichts bis zur Mitte des 16. Jahrhunderts*. 1882. Gotha.
- Wiegand, Herbert Ernst. 1984. Prinzipien und Methoden historischer Lexikographie. I: Besch/Reichmann/Sonderegger (udg.): 557-620.

## Lexikografiska principer för alfabetisk filering med dator

Den alfabetiska ordningen har sedan århundraden varit vår viktigaste nyckel för att komma åt information via lexikaliska register. När vi lämnar det manuella inordnandet och övergår till automatiska datarutiner, upptäcker vi snart att den traditionella alfabetiseringens principer är långt ifrån så triviala och entydiga att de enkelt kan formaliseras. Detta bidrag försöker ge en liten inblick i alfabetiseringens huvudproblem och beskriver en lexikografiskt baserad modell som kan implementeras i våra datorsystem.

### Vad är lexikografisk alfabetisering?

Frågan kan naturligtvis enkelt besvaras med att begreppet avser *ordnandet av ett antal stickord eller flerordiga uppslagsformer i enlighet med alfabetets bokstavsordning*. Detta låter ju mycket enkelt och klart, men tyvärr visar det sig vid närmare eftertanke finnas en mängd delproblem dolda i denna fråga. Låt mig därför ta det hela från grunden och göra en snabb översikt över några av de viktigaste problemen vid alfabetiskt ordnande.

### Vad är det egentliga syftet med en alfabetisk ordning?

Huvudsyftet är förstås att läsaren/användaren så snabbt som möjligt skall kunna finna den information som han eller hon söker efter. Sedan flera hundra år har man utnyttjat den inlärd, fasta ordningsföljd som ges av bokstäverna i alfabetet. Den hävdvunna, alfabetiska följden varierar emellertid en hel del från språk till språk och har också varierat genom tiderna — vi kan exempelvis påminna oss att runraderna, futharkerna, ger en helt annan ordning än dagens alfabet. Bokstavsramsans har en mnemoteknisk funktion, men den återspeglar också ofta en del av det aktuella språkets (morfo)fonematiska särdrag. Vad vi förknippar med en korrekt bokstavsordning är dock, nota bene, inte alltid en linjär följd av bokstäver. (Jfr tabell 1.)

Det bör betonas att en alfabetisk ordningsföljds främsta syfte är att oinitierade användare — utan några särskilda anvisningar — snabbt skall kunna finna den information de söker, dvs. den skall finnas på en intuitivt naturlig plats. Det är just detta syfte som bör vara vägledande för lexikografer och andra som upprättat alfabetiska förteckningar.

### Var finns problemen?

Så länge det gäller att alfabetiskt ordna enskilda ord med enbart små bokstäver, utan accenter

**bokstavsordning**

Ordningsföljden i de svenska, isländska, tyska och spanska alfabeterna.

Svenska	Isländska	Tyska	Spanska
a, à	a	a, ä	a, á
b	á	b	b
c	b	c	c
d	(c)	d	ch
e, é	d	e	d
f	ð	f	e, é
g	e	g	f
h	é	h	g
i	f	i	h
j	g	j	i, í
k	h	k	j
l	i	l	k
m	í	m	l
n	j	n	ll
o	k	o, ö	m
p	l	p	n
q	m	q	ñ, ó
r	n	r	o, ó
s	o	s (ß=ss)	p
t	ó	t	q
u	p	u, ü	r
v, w	(q)	v	s
x	r	w	t
y, ü	s	x	u, ú, ü
z	t	y	v
å	u	z	w
ä	ú		x
ö	v		y
	(w)		z
	x		
	y		
	ý		
	z		
	Ʒ		
	æ		
	ö		

Tabell 1. Ur Nationalencyklopedin, artikeln *bokstavsordning*.

eller specialtecken av något slag, så är det inga större problem. Komplikationerna uppkommer framförallt vid *alfabetisering av större ordböcker och uppslagsverk, bibliografiska kataloger, personnamnsregister, tekniska termlistor och konkordanser och ordindex till stora, autentiska textmaterial*. Hur skall man t.ex. behandla *skillnaden mellan gemena bokstäver och versaler, diakritiska tecken* (som i ç, é, ł och ñ), *ligaturer* (som æ och œ), *digrafer* (som spanskans *ch* eller holländskans *ij*), *bokstavsvarianter* (som *v-w* och *y-ü* i svenskan), *icke-latinska skrivtecken* (t.ex. grekiska bokstäver i vetenskaplig text), *logogram* (&, %, §, etc.), *förkortningar, siffertal* (arabiska och romerska), *skiljetecken, symboler* osv.? Var kan man

exempelvis förvänta sig att en uppslagsboks läsare söker artiklarna *1,3,5-triazin*, *Henri de Toulouse-Lautrec*, *Le Havre*, *'s-Gravenhage*, *Pingvellir*, *Lübeck*,  *$\alpha$ -strålning*, *Karl X Gustav*, *SJ*, *&*, *o.s.v.*, *Cæsar*, *Ærø* och *1984*? Hur skall de nämnda typerna prioriteras sinsemellan och hur skall varje grupp vara rangordnad internt; vilken inbördes ordning är t.ex. mest logisk att ha mellan de olika diakritiska tecknen?

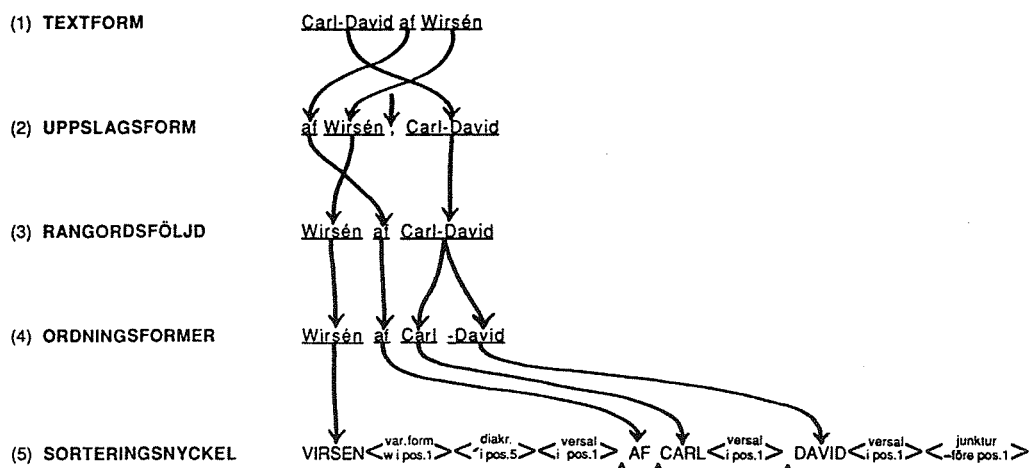
Ett för användaren ofta mycket påtagligt problem gäller också om alfabetiseringen av flerordiga uppslagsformer skall ske *ord-för-ord* eller *tecken-för-tecken* (så att ordgränserna negligeras); det är klart förvirrande när olika principer tillämpas i likartade sammanhang. Skillnaden är förstås särskilt märkbar då uppslagsformens inledande ord är kort. Vid ordvis alfabetisering är det därför t.ex. mindre lämpligt att skriva akronymer med ordmellanrum mellan bokstäverna. Principiellt är en alfabetisering ord-för-ord att föredra om bruket av sär-resp. sammanskrivning är relativt stabilt, som t.ex. i de nordiska språken. Om det däremot vacklar eller varierar fritt kan tecken-för-tecken-principen vara lämpligare. (Det kan dock noteras att t.o.m. Encyclopedia Britannica nyss övergått från sortering tecken-för-tecken till ord-för-ord.)

### Krav på algoritmer för datoriserad alfabetisering

Våra grundläggande krav på principer för att åstadkomma en alfabetisk ordningsföljd är att *ordningen skall bli förutsägbar och entydig*. Detta är särskilt betydelsefullt vid datorbaserad sortering. Kravet innebär att vi måste kunna vara säkra på att få exakt samma ordningsföljd om samma material alfabetiseras på nytt och att det inte blir några komplikationer om vi samsorterar alfabetiserade material från olika källor. Omfattande dataregister för informationssökning och urval ställer också stora krav på explicita och entydiga fileringsmetoder och användaranpassade presentationsformer. *All information som finns i uppslagsordens grafiska form måste därför kunna utnyttjas*, om vi skall kunna tillfredsställa höga lexikografiska krav på stringens och konsekvens. Detta har tyvärr de hittillsvarande standarderna inte klarat av. Jag skall här i korthet beskriva hur en — i tillämpliga delar — datorimplementerbar modell för alfabetisering kan vara uppbyggd.

### Föreslagen alfabetiseringsmodell

Med utgångspunkt i nedanstående illustration (figur 1) skall jag försöka belysa modellens huvudmoment. Vi utgår från de enheter som vi vill använda som sökbara referenser och som skall ordnas i en intuitivt korrekt, alfabetisk följd. De är i sin ursprungliga form ord eller ordförbindelser, så som de faktiskt återges i löpande text (med normal ordföljd). Denna s.k. *textform* (betecknad som nivå 1 i figur 1) är dock inte alltid den vi är vana att använda som *uppslagsform* (2). En viktig grupp där uppslagsformen traditionellt skiljer sig från textformen är personnamnen. I Sverige och flertalet europeiska länder skrivs i sökregister nutida personnamn med *inverterad följd*, dvs. med efternamn, kommatecken och därefter förnamn. Medeltida och äldre personnamn anges dock i *rak följd*, i likhet med bruket på Island (vad beträffar islänningars namn, ofta till skillnad från utlänningars). I vissa fall förekommer det också att andra, flerordiga namn får inverterad ordföljd (som t.ex. typen *Skandinavien*, *Brand- och olycksfallsförsäkringsaktiebolaget*).



Figur 1

En invertering av namnformen är dock inte tillräcklig i alla fall, exempelvis när det gäller hanteringen av namnprefix (*af, von* etc.). I bibliografiska förteckningar behandlas namnprefixen enligt mycket komplicerade regler. Exempelvis brukar prefixet *de* negligeras i namn på danskar, norrmän, holländare, flamländare, fransmän, medeltida italienare, spanjorer, portugiser, m.fl., medan det blir sorteringsgrundande för exempelvis svenskar, engelsmän och nutida italienare. Extra problem uppträder naturligtvis för personer som bytt nationalitet. Så här komplicerade regler är absolut inte användarvänliga — de måste klart förenklas, kanske t.o.m. så mycket att samtliga personnamnsprefix endast blir sekundärt sorteringsgrundande.

Bibliotekens katalogiseringsregler ger oss också ett annat exempel på hur delar av uppslagsformen (t.ex. en boktitel) särbehandlas: inledande prepositioner och artiklar och bestämningar av typen *Svenska, Statens, Allmänna, Aktiebolaget* etc. brukar negligeras; eventuellt får de bilda sista sorteringsgrund. (Effekten blir då i princip likvärdig med en invertering av uppslagsformen.)

För att nämna ytterligare en anmärkningsvärd bibliografisk konvention kan vi tänka på att tecknet & oftast insorteras som om det vore utskrivet som *och, and, und, et, og* etc., alltefter titelns språk.

Det är således befogat att urskilja vad som brukar kallas en *rangordsföljd* (3), dvs. de element (ord) som den alfabetiska sorteringen skall ta hänsyn till, i en preciserad, inbördes ordning; alltså inte nödvändigtvis varje ord från det första till det sista.

För att undvika sorteringsfel som beror på den mänskliga faktorn bör man vara mycket restriktiv med manuella ingrepp i underlaget för sorteringsnycklarna. Hela sorteringsförloppet efter etablerandet av rangordsföljden kan ske helt automatiskt, algoritmiskt — inga manuella ändringar bör därför tillåtas sedan rangordsföljden fastställts; *all information som sorteringen grundar sig på bör vara en entydig återspeglning av uppslagsordens grafematiska form.*

Rangordsföljden är emellertid inte omedelbart användbar som underlag för genereringen av sorteringsnyckeln. Vi måste precisera de explicita teckenföljder som skall

jämföras vid sorteringen. Detta sker i den s.k. *ordningsformen* (4). Här omskrives exempelvis ligaturen *æ* till *oe*, *β* blir *ss* och *ch* blir i spanska material jämställt med ett enda tecken (en kod), jfr tabellen.

I somliga tillämpningar omkodar man på detta stadium arabiska siffertal till en rent alfabetisk form och romerska ordningstal till ett numeriskt värde (för att få korrekt ordning på bl.a. regentföljder).

För vissa applikationer ersätts nu de ortografiskt korrekta ordningsformerna med fonetiska former eller kanoniska normalformer, som sedan får bilda underlag för sorteringen. (Jfr telefonkatalogens normaliserade efternamnsortering, som samsorterar *Carlson*, *Carlsson*, *Karlson*, *Karlsson* etc.)

Vid *finalalfabetisk sortering*, slutligen, vändes ordningsformernas teckenföljd baklänges.

*Sorteringsnyckeln* (5) skapas därefter enligt den s.k. *stavningsprincipen*, dvs. sorteringsnyckeln bygger strikt på (de eventuellt modifierade) ordningsformernas stavning.

I korthet kan följande principer fastställas för hur sorteringen bör fungera och därmed för hur sorteringsnycklarna bör vara konstruerade:

Sorteringen kan ske *ord-för-ord* eller *tecken-för-tecken* (i vilket fall ordmellanrummen negligeras). Trots att skillnaden mellan dessa principer, som nämnts, ofta är mycket påtaglig, finns ingen universellt accepterad konvention eller standard i detta avseende. Svenska telefonkataloger och bibliotekskataloger tillämpar exempelvis *ord-för-ord*-principen medan uppslagsverk traditionellt tillämpat *tecken-för-tecken*-principen. (Enligt den förra sorteras således *allmän väg* före *allmänna* och *Svensk uppslagsbok* före *Svenska Akademien*, enligt den senare metoden tvärtom.) För (svenskt) lexikografiskt bruk förordas att sorteringen sker *ord-för-ord*.

Vid *rent alfabetisk sortering* negligeras till att *börja med alla icke-alfabetiska tecken; vid ordlikhet får övriga tecken* (siffror och skiljetecken) *fälla utslaget*. (Exempelvis hamnar då *1,3,5-triazin* direkt efter *triazin*.) Vid *alfanumerisk sortering* behandlas siffrorna däremot som jämbördiga med bokstäverna i sorteringshänseende; *1,3,5-triazin* hamnar då bland sifferorden.

Jämförelsen av de alfabetiska tecknen görs i varje ord på fyra nivåer:

1. Först jämföres orden i *normaliserad* form, där enbart motsvarande alfabetiska grundtecken och logogram beaktas; vi får då exempelvis följden *& a Abel* ....
2. Vid (ord)likhet tas därefter hänsyn till ev. *varianttecken* i orden: *Lybeck Lübeck ... tvist twist twista twistemål ...* (enl. svensk konvention).
3. För det tredje beaktas eventuella *diakritiska tecken*: *cote coté côte côtelé* ....
4. Slutligen ges bokstäver som utmärks med *versal* lägre prioritet än motsvarande gemen: *ma mA Ma MA* ....

Jämförelsen bör alltså normalt ske *ord-för-ord* samt från vänster till höger, så att ord som börjar på ett likartat sätt hamnar intill varandra. (Vi får då exempelvis ordningen *in in- -in -in- ...*; jfr även 3 och 4 ovan.)

Den modell för alfabetisk filering som jag här kortfattat presenterat finns utförligare beskriven



i Gavare 1988. HÄri finns också utförliga anvisningar för modellens implementering i datorprogram. (Modellen är underlag för Teknisk norm (nr 34) för svenska statsförvaltningen. Den diskuteras för närvarande även inom europeiska och internationella standardiseringsorgan.)

### **Litteratur**

Gavare, Rolf. 1988. Alphabetical Ordering in a Lexicological Perspective. I: *Studies in Computer-Aided Lexicology*: 63-102. Stockholm.

## Proprium og leksikografi

Forfattere drøfter kva plass proprium bør ha i vanlege ordbokverk. Han meiner at ei ordbok som tek sikte på å vera fulldekkjande, også bør ha med proprium. Dei fleste ordbøkene har berre unntaksvis døme frå denne kategorien, og når dei har det, er det ikkje alltid gjeve klare grunnar for utvalet. Dei to viktigaste innvendingane mot å ta med proprium synest vera at materialet er stort og uhandterleg, og at semantikken er innfløkt. Forfattere drøfter dei metodiske vanskanene med å skilja mellom proprium og appellativ. Ei anna sentral problemstilling er korleis ein skal avgjera om opphavlege proprium har gått over til appellativ, t.d. *ford* m. av *Ford*, og i kva omfang slike ord skal takast med i ordbøkene. Han kjem vidare inn på terminologiproblemet i samband med at termen *namn* er nytta om ord i appellativisk funksjon, t.d. *brødnamn* for *brødnemne*. Forfattere konkluderer med at leksikografar og namneganskarar bør drøfta inngåande korleis proprium skal handsamast i ordbokssamanheng.

### 1. Innleiing

Når eg no ynskjer å ta temaet *proprium og leksikografi* opp til drøfting, er det av di det representerer eit fagteoretisk grensefelt for namnegranskarane like mykje som for leksikografane. At mi eiga avdeling, Avdeling for namnegransking, har dagleg kontakt med Avdeling for leksikografi, gjer ikke emnet mindre aktuelt. Dei to avdelingane har sidan 1944 halde hus saman, rett nok i eit omskifteleg tilvære. I seinare år har samværet vorte fagleg utdjupa, ikkje minst gjennom regelmessige husseminar. Ein del av synspunkta i dette innlegget vart fyrst lagde fram på eit slikt husseminar i 1986. Dei tilsette ved avdelingane har òg utnytta samlingane til kvarandre. Særleg har setelarkivet til *Norsk ordbok* vore nyttig i namnegranskinga, og ordboksfolka har på si side kunna leita fram ordtilfang i stadnamnsamlingane.

Kvar for seg har dei to avdelingane teke opp fagteoretiske spørsmål, men det har vore lite diskusjon om kvar namnegransking sluttar og leksikografi tek til - og omvendt. Dette er eit viktig prinsipielt spørsmål. Dessutan har det praktiske konsekvensar for avgrensinga av tilfanget i ordbøkene. Eit avsnitt i innleiinga til *The Oxford English Dictionary* (b. 1:XXVII) gjev ein peikepinn på problemfeltet:

In addition to, and behind, the common vocabulary, in all its diverging lines, lies an infinite number of *Proper* or merely *denotative* names, outside the province of lexicography, yet touching it in thousands of points, at which these names, and still more the adjectives and verbs formed upon them, acquire more or less of connotative value. Here also limits more or less arbitrary must be assumed.

I si velkjende handbok i leksikografi seier Bo Svensén at ein ikkje skal bannlysa personnamn og geografiske namn i språklege ordbøker. "Namn tillhör språket, och deras formella egenskaper (stavning, uttal, böjning etc.) hör i högsta grad hemma i en ordbok" (Svensén 1987:38). Like fullt er det lett å slå fast at proprium spelar ei underordna rolle i dei fleste ordbøkene. Viktigaste grunnen til dette er venteleg den spesielle semantiske funksjonen til

proprium (i tillegg til den store mengda av namn som finst). Fleire granskarar hevdar at proprium ikkje har tyding (intensjon) i ein vanleg kontekst. Klaus Hilgemann (1974:372) seier det slik: "Während die Bedeutung von Appellativen als Definition in einem Lexikon erscheinen kann, lässt sich die Bedeutung eines Eigennamens nicht lexikalisch erfassen". Han meiner at eit proprium ikkje har noko allment, beskrivande innhald (tyding) ut over den funksjonen det har som namn. Andre meiner at som språklege teikn må også namn ha ei tyding ut over referansen (altså ikkje berre denotasjon, men også konnotasjon).

Trass i den semantiske ulikskapen mellom proprium og andre språklege storleikar må ein godta at dei er *ord* i språket. Dei er ei undergruppe av substantiva og må beskrivast som andre ordklassar. Eg ser såleis ei sjølvmotseiing i nærast programmessige titlar som "Name into Word. Names That Have Become Common Property" (Partridge 1949) og *Fra ord til navn* (Tveitane 1983). Her kan eg òg ta med Frederic G. Cassidy (1984:402), som hevdar at "A name is not a word. Words designate classes; names places, or things". Men dersom ei ordbok har som mål å samla alt ordtilfang i eit språk, høyrer også namna med. Dei praktiske vanskane med å samla og ordna alt namnetilfang, ta stilling til homonymi, variantar osv., let eg liggja her.

I ei ordbok vil det vera den historiske (etymologiske) tydinga av eit namn som er mest aktuell - i tillegg til opplysningar om kva det er namn på. Dessutan kan det vera aktuelt å ta stilling til kva motiv som har lege til grunn for det einskilde namnet då det vart til (namnesemantisk tyding), altså om t.d. *Hesthammar* tyder 'formasjon som liknar ein hest', 'stup der ein hest har falle utfor', 'ufs der dei kunne ta seg fram med hest' e.l.

Om ein no likevel vel å halda proprium utanfor - som dei fleste ordbøkene gjer - må ein føreta ein grensegang mellom namn på den eine sida og anna ordtilfang på den andre. Den tyske leksikografen Dieter Nerius (1985:295) peikar på ei "Grenzzone" der nemningar for planter, dyr, sjukdomar, hendingar, produkt o.l. har klassifiserande funksjon og fell i gruppa for appellativ, medan ord for t.d. organisasjonar og institusjonar har identifiserande funksjon og er å rekna for proprium. Etter Nerius høyrer den fyrste gruppa heime i ordbøkene, men ikkje den andre, bortsett frå proprium som har fått appellativisk funksjon. Eit slikt appellativisert proprium er t.d. *kneipp* m.

Eg kjem attende til denne ordtypen nedanfor. Eit spørsmål som meir indirekte er relevant for temaet mitt, er den terminologiske bruken av *namn* i staden for *nemne* (bokmål/dansk *betegnelse*, svensk *beteckning*), t.d. *plantenamn*, *brødnamn*. Også dette vil eg koma attende til.

## 2. Eit klassifikasjonsproblem: Proprium - appellativ

Det finst fleire kriterium på kva som skil proprium frå appellativ. Den reint grafiske skilnaden kjem fram i bruken av stor førebokstav i proprium. Det gjeld t.d. ikkje tysk, som har stor førebokstav i alle substantiv. Proprium har vore definert på ulike måtar i høve til appellativ, m.a. ved at eit namn ikkje vert bøygd i kjønn og tal, at det alltid er bestemt (syntaktisk), at det er individualiserande, men ikkje klassifiserande, at det er tydingslaust. Det har vist seg vanskeleg å få desse kriteria til å fungera som vasstette skott mellom dei to kategoriane.

Eg har i ein annan samanheng (Helleland 1987) drøft propriets sin status i høve til appellativ, og den fylgjande diskusjonen byggjer på framstillinga der. Dei aller fleste namna synest ha oppstått av ikkje-proprialt ordmateriale. Ein kan sjå namnelaginga som resultat av ein "proprialiseringsprosess". Men om denne prosessen har skjedd over ein fase med "halv-appellativisk" funksjon av t.d. eit opphavleg appellativ, eller om propriumsfunksjonen har

oppstått individuelt og spontant, er det ei viss usemje om. Andersson (1973:152) talar om "gränszonen mellan egennamn och appellativ; någon skarp gräns kan inte dras mellan de två substantivkategorierna".

Sjølv meiner eg at anten fungerer eit språkleg uttrykk som eit proprium, eller så gjer det det ikkje. Etter mitt syn er det difor gale å snakka om "grader" av proprium. Det er konteksten, dvs. språkbrukaren, som avgjer funksjonen. Same språkbrukar kan visseleg bruka eitt og same språklege uttrykk med ulik funksjon, men ikkje samstundes. At det i mange høve kan vera ugreitt å avgjera kva for kategori det er tale om, ser eg på som eit analyseproblem.

### 3. Appellativ av proprium

Ei gruppe ord som kan valda problem for ordboksredaktøren, er dei som har oppstått av proprium. Problemet her er for det fyrste å avgjera om eit ord er appellativisk, og for det andre om det i så fall bør koma med i den aktuelle ordboka.

Innbyggjarnemna er illustrerande i så måte. Dei fleste er avleidde av stadnamn, altså proprium, men ut frå sin funksjon er dei appellativ og dermed i prinsippet kvalifiserte til å koma med i ordbokssamanheng (eg kjem attende til terminologien under bolk 4). Dette er ein talrik ordflokk etter som kvart bustadnamn kan gje opphav til eit nemne for innbyggjar. Men av dei ca. 50 000 gardsnamna i landet (rett nok mange likelydande) er det berre ein del som har resultert i allment kjende innbyggjarnemne. I ordbøkene er det truleg berre nemne laga til distriktsnamn og kommunenamn, mogeleg òg soknenamn og tettstadnamn, som det er aktuelt å ta med. t.d. *mandalitt*, *sogning*, *tunesokning*, *tyssedalning*. Eit utval av desse står elles i *Norsk stadnamnleksikon*.

Ein annan type propriumsrelaterte appellativ er slike som er samansette av eit proprium pluss eit appellativ, t.d. *vestlandsbonde*, *oslodame*, *odelstingsproposisjon*, *sigdalsfrukost*, *eidfjordmål*, *geilomøte* (ei tidlegare årviss forsamling av kristne organisasjonar), *holmenkollrenn*. Av desse er dei fire fyrste førte opp i t.d. *Nynorskordboka*. Når dei tre siste ikkje har kome med, er det sjølv sagt eit val redaktørane har gjort. Prinsipielt er det ikkje nokon skilnad på dei - dei fungerer alle som appellativiske samansetningar. I ordbøkene synest kravet til opptak på orda å vera at dei skal ha ein viss allmenn bruksfrekvens, noko som ofte må vera vanskeleg å avgjera. Somme ord har berre forbigåande aktualitet, t.d. er *sigdalsfrukost*, *geilomøte* ord for alt no meir eller mindre historiske tilhøve, og då har dei straks lågare prioritet. Dersom redaktørane hadde hatt tilgjenge til representative frekvensordlister for norsk, ville dei hatt eit sikrere grunnlag for utveljinga.

Samansetningar av det ovannemnde slaget vert ofte skrivne med stor førebokstav av stadnamnet pluss bindestrek, altså *Sigals-frukost*, *Eidfjord-mål*, *Geilo-møte*, *Holmenkoll-renn*. Ein slik skrivemåte er vanleg så lenge ikkje samansetningane vert oppfatta som "reine" appellativ. Det gjeld framfor alt dei mange ad hoc-lagingane som florerer i moderne avisspråk. Men denne grafiske ulikskapen endrar ikkje på den grammatiske strukturen. Det dreiar seg om appellativ der stadnamnet fungerer som adjektivisk underledd til hovudleddet (*frukost*, *mål*, *møte*, *renn*). Rett nok kan dei same språklege storleikane ha funksjon av proprium, t.d. *Geilomøtet 1989* som namn på ei viss tilskiping, men ein slik funksjon må avgjerast ut frå konteksten.

Ei anna stendig aukande gruppe appellativ er dei som har oppstått direkte av proprium. Til dei allment kjende og fagleg uproblematiske høyrer ord som *disel* m., *quisling* m., *røntgen* m., *volt* m., *watt* m. Ingen tenkjer å skriva desse med stor førebokstav inne i ei setning.

Mindre sjølvstakt kan det synast med liten førebokstav i produktord som *ford*, *walkman* (lommeradio med øyretelofonar), *brøyt* (gravemaskin). Men også desse fungerer som appellativ. Ein kan t.d. seia og skriva "forden kolliderte". Derimot bør ein rekna med *Ford* som namn på "prototypen" av dette bilmerket - og av den enkelte modellen med sine spesifikasjonar (overført frå namnet på grunnleggjaren *Henry Ford*). Ei samling på ikring 3000 engelskspråklege appellativiske nemningar med opphav i proprium er lagt fram av Eric Partridge (1949).

Ord for brød kan vera eit godt utgangspunkt for å drøfta appellativiseringsprosessen ved produktamn. Det mest kjende dømet er det ovannemde ordet *kneipp* (*kneippbrød*), som går attende på den tyske lækjaren og presten Sebastian Kneipp (1821-97), kjend for sine helsekurar og brøddoppskrifter. (At *Kneipp* som slektsnamn reflekterer nedertysk *knîp* 'skomakarkniv', er interessant nok, men uviktig i denne samanhengen.) Andre døme på personnamn som har gjeve opphav til brødnemne, er *Graham*, *von Krogh*, *Marianne Rykke*. I somme høve ligg stadnamn til grunn for brødnemne, t.d. *Alaska*, *Bergen*, *Beito*, *Bislett*. I samband med ei gransking av ord for brød (Helleland 1990) bad eg ei gruppe studentar og tilsette ved Universitetet i Oslo om å setja stor eller liten førebokstav på ca. 60 ord for brød i ein kontekst der det aktuelle ordet ikkje stod fyrst i ei setning. Svara varierte sterkt, frå mest berre stor førebokstav til mest berre liten førebokstav. Men det var ein viss dominans for stor førebokstav når brøddordet gjekk attende på namn som er lette å identifisera, t.d. "*Alaska*(brød)". Konklusjonen av dette er, svært forenkla, at folk oppfattar namnekarakteren ved slike ord på ulik måte. Typen *kneipp* er likevel så innarbeidd som nemne at få oppfattar det som eit proprium.

Sjølv om altså ord for brød må reknast til appellativa, har dei som andre produktord propriale uttrykk som base. Til og med gjennomslitige samansetningar som *grilloff*, *ferdabrød* representerer etter mitt syn ei opphavleg namnelaging i og med at dei må ha vore nytta om ei nyskaping. *Ferdabrød* vart såleis namngjeve i samband med ein brødkonkurranse på Island for ein del år sidan (oppløyst av Baker- og konditormestrenes landsforening). Den tyske namnegranskaren Dieter Berger (1976:377) har argumentert for at ord for produkt (varemerke) er proprium. "Sprachlich gesehen geht es auf einen Namengebungsakt zurück, der die Einmaligkeit der Erfindung oder der Herstellungsweise betonen will". Berger tenkjer då på "prototypen", men han aksepterer at når det finst fleire eksemplar av produktet, må det språklege uttrykket for det klassifiserast som appellativ.

Det teoretisk problematiske punktet er når namnet på ein prototype (*Ford*, *Kneipp*, *Ferdabrød*) går over til å fungera som appellativ. Dersom me tek for oss *dovrebrød*, laga til stadnamnet *Dovre*, så har den eller dei som har valt dette som produktamn, på eit eller anna stadium gått ut frå *\*Dovrebrød* som namn på produktideen og prototypen. Fyrst når det vert produsert fleire eksemplar under dette språklege uttrykket, kan ein seia det har vorte eit appellativ. Men det vil gjerne gå ei viss tid før ordet er så godt innarbeidd i språket at det vert aktuelt å ta det opp av ordboksredaksjonane. Eg kan nemna at av dei vel 300 brødnemna eg har drøft i den nemnde artikkelen, har eg funne berre ca. 15 i *Nynorskordboka*.

Dersom ordboksredaktørane tek konsekvensen av den svære tilveksten av appellativiske lagningar av opphavlege proprium, vil talet på oppslagsord måtta auka sterkt. Det gjeld ikkje berre dei allment kjende nemna av typen *kneipp*, *walkman*, *ford*, men òg det breie spekteret av ord for vareslag som *pepsi* (*Pepsi Cola*), *omo* (*Omo*), *pepsodent* (*Pepsodent*), *sunlight* (*Sunlight*). Her kan ein òg ta med dei mange bokstavorda av typen *nyco* (*Nyco*), laga til firmanamnet *Nygård og Co* (sic!) (jf. Grytdal 1937:192). Eit juridisk problem er at mange av desse orda er varemerke, og det har kome reaksjonar frå einskilde produsentar på at

ordbøkene ikkje skil mellom "frie" ord og varemerke. Såleis har Norsk Hydro reagert overfor Norsk språkråd på at *fullgjødsel* er ført opp utan opplysning om varemerke (munnleg opplysning av amanuensis Boye Wangensteen).

#### 4. Eit terminologiproblem: Namn - nemne

Sjølv om ordboksredaktøren har greidd å skilja mellom proprium og appellativ, står det att eit terminologisk problem som kan forvirra. Det gjeld *namn* nytta som *nemne* (*nomen communis*). I det daglege ordboksarbeidet representerer dette neppe noko problem, men det har prinsipiell interesse å drøfta spørsmålet. I det namnefaglege miljøet i Norden har det i dei siste åra vore ein del diskusjon omkring terminologien på dette punktet. Såleis har Thorsten Andersson (1983:11) på eit symposium om personnamnterminologi gått inn for at "termen *namn* ... bör begränsas till att gälla *egennamn*". På eit symposium i Umeå i 1988 om plante-namn vart dette spørsmålet på nytt teke opp, og der argumenterte både Sigurd Fries (1990:9ff) og andre sterkt for at ein ut frå omsynet til fagtradisjon og allmenn språkbruk må akseptera *namn*, t.d. "*växtnamn*", brukt i appellativisk funksjon. For same kva språkgranskarane hevdar, vil *namn* framleis vera vanleg term i folkeleg språkbruk om ei rekkje appellativgrupper. Botanikarar og zoologar vil som før operera med *plantenamn* og *dyrenamn* i sine artsoversyn - som bakarane opererer med *brødnamn*. Dessutan er *namn* kortare og lettare enn lange ord som *betegnelse* (dansk, bokmål) og *beteckning* (svensk). Her har nynorsk rett nok den kortare termen *nemne* (i tillegg til *nemning*).

Denne terminologiske ugreia gjer seg særleg gjeldande ved ord for innbyggjar, t.d. *bergensar*, *odding*, *opedøl*, altså ord som i tradisjonell språkbruk vert rekna som *innbyggjarnamn* (i fagspråket stundom *inkolentnamn*). Her får me nemleg i tillegg det problemet eg har peikt på ovanfor med å skilja mellom proprium og appellativ. Ola Stemshaug (1983) har diskutert "termar for nemning av innbyggjar" og kome til at samletermen *innbyggjarnamn* bør nyttast for heile ordgruppa, og at *innbyggjarnemning* bør reserverast for den appellativiske delen av ordgruppa (Stemshaug 1983:153). Det skiljet Stemshaug gjer framlegg om, synest noko uklart og terminologisk ikkje heilt heldig. Det er vel heller ikkje semje om at det språklege uttrykket for ei folkegruppe kan reknast som proprium. Eg er klar over at det finst eit grenseområde der det kan vera vanskeleg å avgjera kva grammatisk kategori uttrykket høyrer til, t.d. den typen som Stemshaug nemner *gardsfolknamn* 'folket på ein gard'. (Etter mitt syn høver *husfolknamn* betre av di det ikkje alltid er tale om folk på ein gard. Evert Melerfors (1983:13) føreslår *husbondenamn* og reknar denne typen som proprium.) Dersom det språklege uttrykket for slike storleikar kjem i klasse med t.d. *tilnamn*, altså at dei er å sjå på som faste, ubøyelege uttrykk for ei persongruppe, er eg samd i at *namn* vil vera ein korrekt term. Men dersom ein ved analysen kjem til at dei likevel er appellativiske uttrykk, vil *gardsfolknemnel-nemning* (eller etter mitt framlegg *husfolknemnel-nemning*) vera rettare termar.

Døme på ei slik "minigruppe" er folket på gardsbruket *Kiken* i Ullensvang. Dei vert omtala som *Kikanel/kikane*, *Kikafolket/kikafolket* eller *Kika* og *deilkika* og *dei* (med stor eller liten førebokstav etter kva kategori ein vel å sjå uttrykka). Eit argument for at me her har med proprium å gjera, er at typen *Kikanel/Kikafolket* vanlegvis førekjem ubøygde. Men ein kan òg seia at "*ein kike* var òg med", "det kom *kikafolk* etter vegen" kan tenkjast. I dei to siste døma opptrer orda i appellativisk funksjon. *Kikanel/kikane*, *Kikafolket/kikafolket* er altså homonyme storleikar (jf. Dalberg 1985) der den propriale motsett den appellativiske funksjonen er avhengig av konteksten. Sameleis kan ein omtala ein person frå *Opedal* med det appellativiske

uttrykket (innbyggjarnemnet) *opedølen*. I ein annan kontekst kan *Opedølen* vera fast tilnamn på ein person frå Opedal, sameleis som *Kiken* kan vera tilnamn på ein person frå gardsbruket *Kiken*.

## 5. Proprium i ordbøker

I artikkelen "Om historisk lexicografi av typ SAOB" gjev Hans Jonsson (1983:127ff) eit oversyn over korleis *Svenska akademis ordbok* og andre ordboksverk er redigerte. Jonsson skil mellom 1) samtidsordbøker t.d. *Nudansk Ordbog*, 2) eldrespråksordbøker, t.d. *Ordbog til det ældre danske sprog*, og 3) samtids-eldrespråksordbøker, t.d. *SAOB*. Den siste typen tek m.a. opp eldre og yngre former og tydingar av like eller nærskylde ord. Jonsson kjem ikkje inn på avgrensinga i høve til proprium, men Sture Hast (1983:188) seier at "Namn utan användning som appellativer (personnamn, ortnamn, marknamn) medtas ej". Derimot vert namn med funksjon som appellativ og samansetjingar og avleiingar til namn tekne med.

*Ordbog over det danske sprog* drøfter i eit kort avsnitt (b. 1:XXV) "egennavne, egentlige nomina propria" og meiner at dette ordstoffet "der i og for sig har en overordentlig interesse, baade sprogligt og kulturhistorisk, er holdt udenfor de fleste modersmaalsordbøger, selv de meget store". Sjølv har dette verket med berre poetiske namn og appellativ av opphavlege namn.

Ei moderne ordbok med særleg rikholdig propriumstilfang er *Nudansk Ordbog*. Her er mange både stadnamn og personnamn førte opp med språkhistoriske forklaringar. Denne ordboka gjev såleis, rett nok med sitt moderate omfang på to ikkje særleg store band, eit representativt tverrsnitt av det mest frekvente ordtilfanget i dansk språk.

Det store norske ordboksverket *Norsk Ordbok* gjev fylgjande korte fråsegn om proprium: "S t a d n a m n og andre sernamn er til vanleg ikkje medtekne i ordbøker av dette slaget, såleis heller ikkje her<sup>1</sup>. Av ibuarnamn er det teki med eit utval, og elles er det i artiklane om dei vanlegaste namnelekkane (som *bu(e)*, *-dæling*, *-døl*, *-døling* osfr.) teki med døme på ibuarnamn med slik sistelekk" (her er altså den tradisjonelle, men fagleg sett uheldige termen "ibuarnamn" nytta). Eittbindsverket *Nynorskordboka* har ikkje med proprium og seier heller ikkje noko om denne kategorien.

I fortalen til *Norsk Ordbog* (Aa) seier Ivar Aasen (s. XV) at han har teke med ein del språkhistorisk viktige ord "som kun høres i Stedsnavne og Personnavne, og hvis Betydning mange kunde ønske at vide" (med tillegg i fotnote om at det gjeld slike som *Ang*, *Lo*, *Vin*, *Gunn*, *Leiv*, *Mund*, *Rid*, *Vald*). Men han understrekar at det gjeld få ord, etter som alle dei som vert brukte i eineståande namn, er forbigått: "Ligeledes er alle virkelige Egennavne (eller Proprier) udeladte, uden for saavidt som de kunde nævnes i Forklaringen ved de Ord, som afledes af saadanne Navne."

Hans Ross har i si *Norsk Ordbog* (Ro) ikkje uttrykt noko særskilt om proprium, bortsett frå at han (s. XV) seier at "Stedsnavnene er i det væsentlige skrevne efter Egnens Udtale, undertiden med nogen Tillempning til den gamle Form", men her er det venteleg tale om stadnamn nytta som lokalisering og ikkje som døme. Han har likevel med ikkje så få oppslag der eit stadnamn ligg til grunn for oppslagsordet, og der stadnamn er nytta som døme på bruken. Stundom kan det vera vanskeleg å avgjera om eit oppslag er heimla som proprium eller appellativ, som når han under *Nyk* m. har forklaringa "steil Bjergknold=Nuk". Under *Nuk* m. i *Tillæg* og *Rettelser* fører han opp "høi Bjergknold. Isl. *hnjúkr* - Helg. Hertil, og til *Nyyk*: "Njeutkjinn", Navn paa en Fjeldtop i Voss (Ævaanga), og "Knyytkjinn", en do. i Ndm.

(Surnadal)." (I kartoteket til Norsk Ordbok er det eit par døme på at *nyk* m. er ført opp som appellativ, men i båe høve med tilvising til visse fjell.)

*Ordbog over det gamle norske Sprog* av Johan Fritzner har prinsipielt ikkje med namn i oppslagsform, men tilvising til stadnamn i døme som *ló* f. "i Bygdenavnet *Lóar*" (b. 2, s. 550), *sygnir* "Mænd fra Landskabet Sogn i Norge, deraf *Sygnafylki*" (b. 3, s. 629). Og om ordet *anгр* m. er fylgjande opplyst: "Fjord, Bugt, forekommer dog kun i Stedsnavne" (b. 1, s. 56). *Gamalnorsk ordbok. Med nynorsk tyding* (1930 og seinare) av Leiv Heggstad har derimot med eit stort tal gamalnorske stad- og personnamn. Under oppslagsordet *rust* f. er det t.d. vist til *í Rustum*. (Men i den reviderte utgåva *Norrøn ordbok v/ Heggstad, Hødnebo og Simensen frå 1975* er oppslaget, vel med grunnlag i etymologiske ordbøker, endra til *Rústum* "dat. i stadn.", altså til eit ord med opphavleg lang vokal. Det er såleis noko misvisande når det her samstundes er referert til nyno. *rust* f. 'holt, liten skog', som har opphavleg kort vokal (ikkje belagt i gamalnorsk).

I ein instruktiv artikkel om namnestoff i ordbøker seier Ebba Hjorth (1979:39) at *Gammeldansk Ordbog* fylgjer prinsippet om at "proprier i funktion som proprier ikke optages i ordbogen, men at sted- og personnavne i funktion som appellativer medtages, at adjektiver afledt af stednavne og sammensætninger med stednavne som førsteled medtages, samt at indbyggernavne er appellativer og selvfølgelig derfor kommer med". Ho meiner (s. 45) at proprialt ordstoff kan auka kunnskapen om ordtilfang som elles er svakt dokumentert. Om eit ord ikkje er belagt som appellativ i perioden for *Gammeldansk Ordbog* (1100-1515), men berre som proprium eller del av eit proprium, bør det takast med om det ut frå anna nordisk språkmateriale er grunn til å tru at det har eksistert, t.d. *\*løn* m. i stadnamnet *Lønholt*, skrivre in *Lønholtæ* 1348.

For dei etymologiske ordbøkene kan stadnamn ha særleg interesse. Alf Torp har i si *Nynorsk etymologisk ordbok* ei rekkje døme på ord som er relaterte til stadnamn. Under *núp* (s. 464) har han t.d. vist til shetlandsk *nup* "'bratt fjeldtop', som stedsnavn". Om *herad* seier Torp at det finst "mest i navn, som Gransherad" (s. 210). Liknande døme vil ein finna fleire av. *Svensk etymologisk ordbok* av Elof Hellquist har med "Ett stort urval av *förnamn* och ett mindre av *ortnamn* och *familjenamn*" (s. VI). I *Wortbildungslehre* nyttar Wolfgang Meid ofte person- og stadnamn til å visa døme på avleiing.

Eg skulle gjerne sjå eit nytt norsk etymologisk oppslagsverk som femnde om alt heimla ordmateriale, òg materiale som berre finst i namn, t.d. eit øynamn som *Huglo* (til same rot som haug, jf. tysk *Hügel*), og eit elvenamn som *Kinso* (til *kinn* f. om bratt fjellside). Her kunne ein tenkja seg namnet i si noverande form som oppslagsord, eller ei konstruert eldre form forsynt med asterisk. Det siste har t.d. *Norske Gaardnavne* gjort med dei mange konstruerte orda som er postulerte i stadnamn.

## 6. Namneleksikon og namnesamlingar

Når det gjeld leksikalske oversyn over *namn*, finst det fleire som tek for seg personnamn (førenamn). *Norsk personnamleksikon* gjev flest språklege opplysningar. *Norsk stadnamleksikon* har forklaringar om dei viktigaste stadnamna i landet. Danmark har eit tilsvarende oppslagsverk i *Dansk Stednavneleksikon*. Dei norske gardsnamna er lagde til rette og forklara i Oluf Rygh: *Norske Gaardnavne*. Ut over dette ligg det føre ein omfattande vitskapleg litteratur der personnamn og stadnamn er drøfte typologisk og semantisk.



Av meir komplette *namnelister* over personnamn (førenamn og slektsnamn) utan tydingsopplysningar nemner eg *Personregisteret* i Statistisk sentralbyrå med fullstendig oversyn over personnamn i bruk. Den tidlegare *Grunnboka*, no ein del av *Grunneigedomss-, adresse- og bygningsregisteret* (GAB) inneheld namn på ein stor del av eigedomane i landet. Statens kartverk har eit databasert register over namn på hovudkartserien i 1:50 000 (N50). Kartverket har som mål å leggja alle stadnamn på offentlege kart inn i eit sentralt databasert register. Namnearkiva i samarbeid med regionale og lokale instansar arbeider med å oppretta ein *norsk stadnamnbase* som skal utgjera eit framtidig grunnlag for gransking og utgjeving av stadnamn. Også andre namnegrupper går inn i samlingane, t.d. tilnamn og husdyrnamn. Namn som til no har vore av mindre interesse for namnearkiva, er slike som gjeld transportmiddel, institusjonar, tiltak, kunstverk og andre produkt. Like fullt dreiar det seg om eit namnegods som spelar ei viktig rolle i daglegspråket vårt. Noko samla register over alle desse namna finst ikkje - her må ein gå til ulike instansar for skaffa seg oversyn. Men eit omfattande leksikon med tydingsopplysningar av alle namn - på line med ikkje-proprialt ordtilfang - er førebels ikkje noko konkret prosjekt.

## 7. Sluttord

Både leksikografen og namnegranskaren ber vel inst inne på ein draum om at alt ord- og namnemateriale skal gjerast tilgjengeleg i oppslagsform og i kontekst. Draumen kan neppe nokon gong realiserast fullt ut, men den maskinelle ordboksteknologien opnar for store perspektiv. I dei framtidige elektroniske ordbøkene kan ein tenkja seg at propria får sitt rom saman med det ikkje-propriale ordtilfanget med tilhøyrande tydingsforklaringar, bøyingsformer og bruksopplysningar. Av ein slik grunndatabase kan det så genererast ordbøker for ulike spesialområde.

Om eit altomfattande leksikon både prinsipielt og praktisk enno i lang tid må vera eit diskusjonstema, er det eit anna felt der samarbeidet mellom ordboksfolk og namnegranskarar bør gje snøgge resultat, og det er i utnyttinga av ord som i dag berre er registrert i namn. Både stadnamn og andre namnegrupper inneheld ord som elles ikkje er belagt. I hovudoppgåver og andre vitenskaplege framstillingar om namn finst det til saman mange slike ord, for ikkje å snakka om materiale som ikkje er vitenskapleg handsama eller ikkje eingong innsamla.

Etter mitt syn burde leksikografar og namnegranskarar koma saman og drøfta dette spørsmålet. For eit verk som *Norsk Ordbok* ville det vera mykje å vinna ved ei slik utviding. At det enno står att mykje registreringsarbeid før den norske ordskatten er gjort tilgjengeleg, bør ikkje vera til hinder for eit slikt samarbeid.

## Notar

1. Magne Myhren, ein av redaktørane i *Norsk ordbok*, opplyser at han i manuskriptet under bokstaven *g* har teke med *grev* n. 'utgravi dokk', med tilvising til Gustav Indrebø: *Norske innsjønamn i Buskerud fylke*, Oslo 1933, s. 64. Etter Indrebø er ikkje ordet kjent utanom stadnamn.

## Litteratur

- Andersson, Thorsten. 1973. Om ortnamn och ortnamnsforskning. I: *Namn och bygd* 61: 148-170. Uppsala.
- Andersson, Thorsten. 1983. Personnamn. Till begreppets avgränsning. I: *Personnamsterminologi. NORNAs åttonde symposium i Lund 10-12 oktober 1981*. NORNA-rapporter 23. Redigerad av Göran Hallberg, Stig Isaksson och Bengt Pamp: 9-22. 1983. Uppsala.
- Cassidy, Frederic G. 1984. The Etiology of Place Naming. I: *Names*. Vol 32, No. 4: 402-406. New York, NY.
- Berger, Dieter. 1976. Zur Abgrenzung der Eigennamen von den Appellativen. I: *Beiträge zur Namenforschung*. Neue Folge 11 (1976): 375-387. Heidelberg.
- Dalberg, Vibeke. 1985. On Homonymy between Proper Name and Appellative. I: *Names*. Vol. 33, No. 3: 127-135. New York, NY.
- Dansk Stednavneleksikon = Jørgensen, Bent. *Dansk Stednavneleksikon*. [1] *Øerne øst for Storebælt*. København. 1981. [2] *Jylland - nordlige del*. København. 1982. [3] *Jylland. Sydlige del. Fyn med omliggende øer*. København. 1983.
- Fries, Sigurd, red. 1990. *Om växtnamn. Tio föredrag vid växtnamnssymposiet i Umeå den 6-7 december 1988*. Nordsvenska. Skrifter utgivna av Institutionen för nordiska språk vid Umeå universitet. Umeå.
- Gamalnorsk ordbok = Heggstad, Leiv. *Gamalnorsk ordbok. Med nynorsk tyding*. Ny umvølt og auka utgåve av "Gamalnorsk ordbok" ved Hægstad og Torp. Uendra opptrykk av 2. utg. (1930). Oslo 1963. Revidert utg.: *Norrøn ordbok*. Ved Finn Hødnebo og Erik Simensen. Oslo. 1975.
- Grytdal, Normann. 1937. Bokstavordene i språket vårt. I: *Maal og Minne*. Norske studier. Utgitt av Bymaalslaget: 187-193. Oslo.
- Hast, Sture. 1983. Om SAOB:s material. I: *Nysvenska studier*. Tidskrift för svensk stil- och språkforskning. 63: 159-192. Uppsala.
- Helleland, Botolv. 1987. Namn og nemne (NN). Eit leksikalsk-onomastisk problem. I: *Namn og Nemne* 4 (1987): 23-32. Øvre Ervik.
- Helleland, Botolv. 1990. Ord på vårt daglege brød. I: *Institutt for namnegransking. Årsmelding 1989*: 77-100. Oslo.
- Hilgemann, Klaus. 1974. Eigennamen und semantische Strukturen. I: *Beiträge zur Namenforschung*. Neue Folge 9 (1974): 371-385. Heidelberg. Også I: *Reader zur Namenkunde 1. Namentheorie*. Utg. av Friedhelm Debus og Wilfried Seibicke: 21-40. Hildesheim-Zürich-New York. 21-40.
- Hjorth, Ebba. 1979. Om udnyttelse af navnestof i ordbøger. Proprier i Gammeldansk Ordbog. I: *Danske studier*. København. 39-49.
- Jonsson, Hans. 1983. Om historisk lexikografi av typ SAOB. I: *Nysvenska studier*. Tidskrift för svensk stil- och språkforskning. 63: 127-158. Uppsala.
- Melefors, Evert. 1983. *Byngen, Smissen och Listar. Inbygggarbeteckningar på Gotland, I. Typologi och ordbildning*. Studier till en svensk ortnamnsatlas utgivna av Harry Ståhl. 13. Uppsala.
- Nerius, Dieter. 1985. Zur Stellung der Eigennamen in Wörterbüchern der deutschen Gegenwartssprache. I: *Symposium on Lexicography II. Proceedings of the Second International Symposium on Lexicography May 16-17, 1984 at the University of Copenhagen*. Edited by Karl Hyldgaard-Jensen and Arne Zettersten. Lexicographica. Series Maior 5: 287-301. Tübingen.

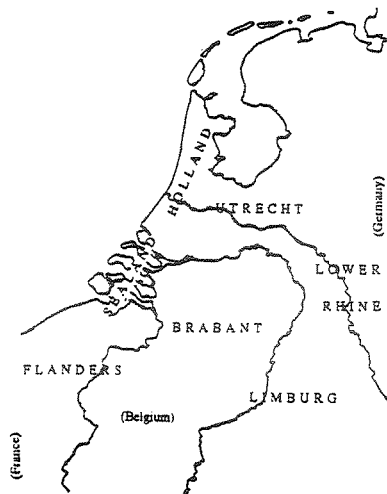
- Norske Gaardnavne = Oluf Rygh: *Norske Gaardnavne. Oplysninger samlede til Brug ved Matrikelens Revision*. Kristiania. 1897-1924.
- Norsk Ordbog (Ro) = Ross, Hans: *Norsk Ordbog. Tillæg til "Norsk Ordbog" af Ivar Aasen. Tillæg og Rettelser*. Christiania 1895.
- Norsk Ordbog (Aa) = Aasen, Ivar: *Norsk Ordbog med dansk Forklaring*. 2. utg. Christiania. 1873.
- Norsk Ordbok = Beito, Olav T. et al. (red.): *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. 1-. Oslo. 1966-.
- Norsk riksmålsordbok = Knudsen, Trygve et al. (red.): *Norsk riksmålsordbok*. Utgitt av Riksmålsvernet. 1-2. Oslo 1937-57.
- Norsk personnamleksikon = Stemshaug, Ola (red.): *Norsk personnamleksikon*. Oslo. 1982.
- Norsk stadnamleksikon = Sandnes, Jørn og Ola Stemshaug (red.): *Norsk stadnamleksikon*. 3. utg. Oslo. 1990.
- Nynorsk ordliste = Hellevik, Alf: *Nynorsk ordliste*. Større utgåve. Oslo. 1990.
- Nynorskordboka 1986 = Hovdenak, Marit et al. (red.): *Definisjons- og rettskrivingsordbok*. Utarbeidd av Avdeling for nynorsk ved Norsk leksikografisk institutt, Universitetet i Oslo, i samarbeid med Norsk språkråd. Oslo.
- Nynorsk etymologisk ordbok = Torp, Alf: *Nynorsk etymologisk ordbok*. Kristiania. 1919.
- Ordbog over det danske sprog = *Ordbog over det danske sprog*. 1-28. Udgivet af Det danske Sprog- og Litteraturselskab. København. 1918-56.
- Ordbog over det gamle norske Sprog = Fritzner, Johan: *Ordbog over det gamle norske Sprog*. I-III. 4. utg. nytt og uforandra opptrykk av 2. utg. (1883-96). Oslo 1973. B. IV. *Rettelser og tillegg*. Ved Finn Hødnebo.
- Partridge, Eric. 1949. *Name into Word. Proper Names That Have Become Common Property. A Discursive Dictionary*. London.
- Stemshaug, Ola. 1983. Termar for nemning av innbyggjar. *Personnamnsterminologi*. NORNA-rapporter 23: 145-157 (sjå Andersson 1983).
- Svensén, Bo. 1987. *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. Stockholm.
- Svensk etymologisk ordbok = Hellquist, Elof: *Svensk etymologisk ordbok*. Lund. 1948.
- Tanums store rettskrivningsordbok = Fossetøl, Bernt og Marius Sandvei: *Tanums store rettskrivningsordbok*. Bokmål. 7. utg. Oslo. 1989.
- The Oxford English Dictionary = *The Oxford English Dictionary. A New English Dictionary Based on Historical Principles*. 1-12. Oxford University Press. London. 1970.
- Tveitane, Mattias. 1983. *Fra ord til navn. Språklig navnegranskning*. En artikkelsamling ved Mattias Tveitane. Bergen-Oslo-Stavanger-Tromsø.
- Wortbildungslehre = Krahe, Hans, og Wolfgang Meid: *Germanische Sprachwissenschaft. III. Wortbildungslehre* (Wolfgang Meid). Band 1218/1218a/1218b. Berlin. 1967.

## The description of names in the *Vroegmiddelnederlands Woordenboek*

The *Vroegmiddelnederlands Woordenboek* (*Dictionary of Early Middle Dutch*) is a period-dictionary, based on historical principles, which will describe the Dutch stock of words recorded in the 13th century. In this dictionary the onomastic material is completely incorporated into the articles. By means of various examples I will show in what way the 13th-century first names, surnames and geographical names are handled in the dictionary and what kind of information they add to it.

### Introduction

At the Institute for Dutch Lexicology in Leiden a team of five editors is working on the *Vroegmiddelnederlands Woordenboek* (*Dictionary of Early Middle Dutch*, henceforth *VMNW*) since 1988. The *VMNW* is a period-dictionary, based on historical principles, which will describe the Dutch stock of words recorded in the 13th century. In that century the change from Latin to the vernacular took place in the administration and the oldest literary products in Middle Dutch date back that far too. The language area involved stretches from Holland, Utrecht and the Lower Rhine in the North to Flanders, Brabant and Limburg in the South.



The source material for the *VMNW* comprises all extant Middle Dutch manuscripts written in the 13th century. This means that not only the charters are included, but also the lexicographical and literary material that has survived from that period. The major part of the material is formed by the *Corpus van Middelnederlandse teksten (tot en met het jaar 1300)*, compiled by Maurits Gysseling. This corpus, usually after its compiler called *Corpus-Gysseling*, consists of two sections. The first section (further referred to as *Corpus I*) contains almost two thousand charters and other official documents, whereas the other section (further referred to as *Corpus II*) contains different kinds of literary texts, varying in length from a few lines to hundreds of pages and in genre from love poems to a translation in verse of Petrus Comestor's *Historia Scholastica*.

The editors' original assignment was to produce a dictionary based solely on the *Corpus-Gysseling*. A small number of 13th-century manuscripts, however, was not included in the *Corpus-Gysseling* for various reasons. For the sake of completeness it has been decided to add this material too. It includes some larger documents, such as the *Glossarium Bernense* (a Latin - Middle Dutch wordlist), the *Diatessaron Leodiense* (a gospel harmony) and *Het oudste goederenregister van Oudenbiezen (1280-1344)* (a register of lended properties). Some other manuscripts, discovered only after publication of the *Corpus-Gysseling*, are an obituary list from Antwerp, some charters from Bruges and a literary fragment of the *Floris*.

All material has been coded with help of the computer, to facilitate description in the *VMNW*. The codes contain a variety of information about date, location and genre of the texts. The words were also provided with information about word class and flexion and with a modern Dutch head word, bringing inflected forms together. Subsequently all texts were put into a database, from which the editors can now retrieve their material by means of head word or code(s). The total number of words, on which the *VMNW* is based, outruns 1,500,000. They will be described in approximately 30,000 articles. It is, of course, a great advantage that the size of the 13th-century stock enables the editors to include all recorded words in the *VMNW* and no selection needs to be made. After the 13th century the number of recorded texts in the vernacular grows explosively, which makes it for practical reasons almost impossible to include all words in a dictionary describing a longer or later period. Then the problem of making a balanced selection comes up.

An important part of the database of Early Middle Dutch words is taken by the onomastic material. Some 12% of the words in the 13th-century texts are names and in articles they will take circa the same percentage. All names will be included in the *VMNW*, in the first place because they appear in the *Corpus-Gysseling* and as such have to be described according to the assignment. Secondly, it is obvious that names have their place in the Early Middle Dutch linguistic system. While studying them sufficient linguistically relevant and interesting facts come up to justify their inclusion in the dictionary. This need not be a surprise: names are linguistic signs with their intrinsic meaning. They are formed after the rules of spelling, morphology and flexion and are - moreover - interesting from a dialect-geographical point of view.

Not only the editors of the *VMNW* support this opinion. An enquiry among a number of potential Dutch and Flemish users of the *VMNW* made clear that a majority of those enquired supported the idea that names are linguistically interesting, and thus thought onomastic material deserved the same treatment in the *VMNW* as other words.

Often in lexicography names have been discriminated, because they were considered less reliable than common nouns, amongst others with regard to their spelling and meaning.

The best chance for a name to appear in a dictionary is usually either to have a lexicalised meaning or to be part of a proverb. This is, for instance, the policy of the *Woordenboek der Nederlandsche Taal* (*Dictionary of the Dutch Language*), which describes the Dutch language from the middle of the 16th century until the beginning of the 20th century. Dictionaries describing earlier language periods tend to be less strict at this point, probably because the later the period, the clearer the distinction between words and names. In the *Middel-nederlandsch Woordenboek* (*Middle Dutch Dictionary*, henceforth *MNW*), for example, produced by the end of the 19th century and describing the stock of words between 1200 and 1500, names have indeed sometimes been included. This regards some (variants of) first names and some words only recorded as surnames. In even fewer cases a place name or place name element is described. It is clear that the description of onomastic material in this dictionary is in no way complete, that the selection of the names seems at will and by chance and that adequate, linguistically relevant information is often lacking. Some examples of this are shown in Schoonheim (1990: 332-333).

In this contribution I will show, by means of 13th-century Dutch examples, in what respects a systematic description of onomastic material can add information to a historical dictionary and in which way this information can be integrated in such a dictionary without losing the distinction between names and other words.

### First names

First or proper names will be extensively treated in the *VMNW*. All recorded variants of a name are entered and described in separate articles. The main emphasis lies on the description of various aspects of the name, amongst others the rise of pet names and augmentatives, Dutch sound-shifts and regional spreading. Under the etymology the original meaning of the name is provided, usually dating back to a Germanic, Greco-Roman or even Hebrew tradition. A linguistic researcher will probably be at least as interested in learning e.g., that the Middle Dutch form *liclais*, only found in the region Holland, is a variant of the name Nicolaas in which the process of distance assimilation has taken place (*n* adapted itself to *l*, so *niclais* became *liclais*), as in knowing that the name Nicolaas originates from the Greek, means 'conqueror of/with the people' and has reached the Low Countries via the Latin language.

Not every variant can for sure be attached to one name. The female first name *marsoete*, for example, can be a pet name to Maria as well as to Margareta. By giving every variant its own entry in the dictionary, the editors can provide information for that particular variant, e.g. about frequency, morphology and localisation. Variants may have been created in a diversity of ways, some of them even regionally bound. The name Nicolaas e.g. has fifteen variants, some of them only found in restricted areas and it is interesting to see how the variants of this name are spread over the regions. A general survey of all recorded variants is presented in what is regarded as the main article on the name. This survey gives one a clear insight into the regional spreading of variants. For the name Nicolaas, for instance, this is done at the entry *nicholaus*, the Middle Dutch form which is closest to the Latin name. At this entry variants as *clais* and *claus*, *liclais* and *niclais* and *cole*, *colin* and *colard* are mentioned and compared mutually. An example of the description of a cluster of first names in the *VMNW* is found in Schoonheim (1990:334-336). This method enables the editors to describe exactly what variety is found and how this variety is spread over the regions. Thus, the user can look up in the *VMNW* the variant he found in his text and will find the

information about this particular form. Cross-references point to the extra (mainly dialect-geographical) information in the main article in the *VMNW* and to other interesting articles. Suffixes of first names, such as *-kin*, *-lin*, *-sin* (all diminutive) or *-en* (augmentative) are treated in separate entries, which enable the editors to give all detailed information there. This saves much space at the other entries and it provides a conveniently arranged survey of the suffix.

Sometimes first names come up from unexpected corners. While describing the words *vlaminc* and *hollander*, which refer to ethnic origin, neither of these words turned out to have been recorded as a common noun, and both appear only in onomastic contexts. Their use as a surname was to be expected, but it was discovered that they have also been used as first names. *Hollander* was found four times as first name, for instance: *hollandre van Reimarswale* (*Corpus I*, page 1454, line 7; Sealand 1290) and *boudine van jiersike hollanders broder* (*Corpus I*, page 1699, line 1-2; Holland, the count's chancellery 1291). The name *vlaminc* was found three times, twice accompanied by a surname: *Vlaming cortals* (*Corpus I*, page 71, line 45; Grauw 1260) and *vlaminc vter haghe* (*Corpus I*, page 2262, line 17; Boechoute 1296), and once in a list of the first names of the magistrates of Boechouterambacht: *jacob. pieter, willem, vlaminc*. (*Corpus I*, page 2262, line 34; Boechoute 1296). Although similar names such as François and Scott are still in use, *vlaminc* and *hollander* are no longer used in the Dutch language area as first names. On enquiry at the Registry Office it turned out that they are not even allowed as first names nowadays. However, the *Woordenboek van voornamen* (*Dictionary of first names*) by Van der Schaar mentions that the name *vlaminc* is still in use in Denmark and Sweden, although slightly changed to *Flemming*, due to *i*-umlaut (Van der Schaar (1984), p. 173) and this 'Scandinavian' name may be given to a child in The Netherlands. Unfortunately, I have found no signs of the use of the word *hollander* as first name after the 13th century.

In the *VMNW* the definition of *hollander* is split in two; under 1 the use as a surname is presented, mainly in the collocation *die hollander* or (with proclitic article) *dollander*. After the references the use as a first name is described under 1.1. Here cross-reference to *vlaminc* is made. It is important to realise that both *vlaminc* and *hollander* are found in onomastic contexts only. If the *VMNW* did not include names, these words would not be described at all, whereas the occurrence as a name logically leads to the conclusion that they must have existed as common nouns too.

The entry *hollander* in the *VMNW*:

## HOLLANDER

znw.m.

### Hollander

[11]

0/11/0 Dordrecht 1284. hollander, dollander (met proclitisch lidwoord); ns -; gs -s. Van het znw. hollant met het suffix -er.

#### I 1. Hollander, iemand uit Holland.

In de verbinding die hollander. Ook met proclitisch lidwoord: dollander. De **Hollander, toenaam van verschillende personen.** || hughe die hollander in monster enen smaltiende. van viif entventich scillinghen. *Corpus I*, p.500, r.21-22 (Holland, graf.kans. 1280-1287). hier over waren yroepen late vanden brieven van aertrike dese dinc te orcondene. bi namen (...) jan

dollandere. hughe dollandere. Ende wouter dollandere. *Corpus I*, p.1701, r.44 - p.1702, r.02 (*Brugge 1292*).

**1.1. Hollander, mannelijke persoonsnaam.** *Zie ook* vlaminc. || Wi Jan van Rinesse. (...) pieter pieters zone. (...) Jan van scenghe. hollandre van Reimarswale. Colin van staple. (...) & andre die van onsen acorde sijn. doen te wetene (...). *Corpus I* p.1453, r.45 - p.1454, r.09 (*Zeeland 1290*). (*In een rekening:*) Item ver. yden hollanders moye .x. soł van haren huse te brekene et xx 5, *Corpus I* p.784, r.12-13 (*Dordrecht 1284*).

## Surnames

A large part of the onomastic material consists of surnames. I here define a surname as any further identification of a person following his first name. Many surnames are place names or patronymics to indicate the roots of the person mentioned. Most surnames, however, are common nouns. Some of these are recorded both in 'normal' use and as a surname. In the description of these words there has been distinguished between the function of the word in ordinary contexts and the function of the word as name. It is not always clear if someone's surname refers to his profession, although mainly this seems to be the case. For the entry *grauwerkere*, for example, the main definition is: 'Fur worker, someone who prepares fur, in particular fur of the Russian squirrel'. (N.B. The first reference mentions someone called *scoreel* 'squirrel', so the man seems to be named after the material he prepares). With regard to the use of *grauwerkere* as a surname, the definition under 1.1 reads: 'Grauwwerker, surname of various persons, in which the distinction between surname and profession name is not always obvious from the context'. Although most persons mentioned under 1.1 may have been fur workers, this is not certain and it is therefore better to keep them apart. The emphasis under 1.1 lies in the fact that their name is 'Grauwwerker', rather than their profession.

The entry *grauwerkere* in the *VMNW*:

**GRAUWERKERE**

znw.m.

grauwerker

[11]

0/11/0 Gent 1270. *graeuwerkere*, *gra(u)werkere*, *grauuerkre*, *grauwarkere*, *ns -*; *gs -s*; *ds -*. *Uit het znw. grau* 'grijs bont' *en het znw. werkere* 'bewerker'.

**I 1. Bontwerker, bewerker van grijs bont, met name dat van de Russische eekhoorn.** || *ian scoreel graeuwerkere*. *Corpus I* p.169, r.37 (*Gent 1270-1290*) N.B. De toenaam *scoreel* betekent 'eekhoorn', het materiaal dat de persoon in kwestie bewerkt, vgl. *Ofra. escurel*. *Item elc grauwerkere sal gheuen van sinen grauwerke* (...) *van der .marc. viiii. ð. Corpus I* p.1337, r.4-6 (*Gent 1288-1301*).

**1.1. Grauwwerker, toenaam van verschillende personen, waarbij het onderscheid tussen beroepsnaam en toenaam niet altijd even duidelijk**



**uit de context blijkt.** || (*Op een dodenlijst:*) Jacop. clais Grauwerkers soene / mhieil (*l. mihiel*) grauwerkere, *Corpus I p.172, r.20-21 (Gent 1270-1290)*. Ook in de verbinding die grauwerkere. || (*Op een lijst van leden van het St.-Jakobsgilde te Gent:*) Niclaus de grauwerkere. *Corpus I p.169, r.5 (Gent 1270)*. (*In een rekening:*) vanden her rasen den grauwerkere .vii tō & viii s & viii ð, *Corpus I p.532, r.15 (Gent 1280)*.

Sometimes a word is recorded in onomastic material only, which implies the meaning has to be given with some reserves. A name can be meant ironically or figuratively; someone who is called *die langhe* ('the tall one') may very well have been rather small. According to the *MNW*, *bouden diserine* ('Baldwin the Iron'), mentioned in three different charters, was given his name because he always rode in harness through the city. However, this surname can as well express a figurative meaning: 'strong as iron; cruel, harsh'. In the case of *clais de iserine* (*Corpus I* page 178, line 36; *Gent 1270-1291*) we have no idea whether to take the surname literally or figuratively.

A noun only recorded as a surname is *crudepenninc*. This word has only been found in Oudenaarde (East-Flanders) and probably referred to one specific family. It is not entered in the *MNW*, but various compounds ending in *-penninc* have been recorded, in which this last part means 'tax on (the earnings from the sale of) ...'. The first part of the compound specifies the subject of the tax, for example *swijnpenninc* 'tax on (the earnings of the sale of) pigs'. It is not unlikely that *crudepenninc* has a similar meaning, like 'tax on (the earnings of the sale of) various herbs'. How a word with such a meaning can become a surname is not quite clear; maybe the tax called *crudepenninc* was collected by this family. Anyway, the fact that *crudepenninc* is only recorded in onomastic contexts does not make it a less interesting word for the dictionary. Although the exact meaning cannot be derived from the word in its own context, it is possible to give a definition with the help of comparable words, so this name, which also as a common noun obviously had a meaning and a function in 13th-century Middle Dutch society, gets its own place in the *VMNW*.

The entry *crudepenninc* in the *VMNW*:

**CRUDEPENNINC**

znw.m.

**Crudepenninc**

[9]

0/9/0 Oudenaarde 1272. *crudepenninc*, *crupenninc*. *ns -;* *gs -s*, *crudepenninx*. *Uit het znw. cruut 'kruiden' en het znw. penninc 'belasting'*.

I 1. **Crudepenninc, belasting op (de inkomsten op de verkoop van) verschillende kruiden.** *Alleen als toenaam van verschillende personen te Oudenaarde aangetroffen.* || *iohannis dicti crudepenninx. Corpus I p.214, r.28 (Oudenaarde 1272)*. *Bachten ser arnouds crudepenninx .i. bunre. ende xxxv. roeden. (land), Corpus I p.1036, r.14 (Oudenaarde 1285)*. *Ende op dat selue erue (...) .vi. s siaers die ghelt gillis crudepenninc, Corpus I p.1584, r.2-3 (Oudenaarde 1291)*.

### Geographical names

An important task of a historical dictionary being to explain words, it goes without saying that a good understanding of historical texts sometimes also requires knowledge of the geography of the time. In the legend of Saint Servatius, dating from circa 1200, a shipping route from Soissons in France to Denmark (Middle Dutch *denemarken*) and Norway (Middle Dutch *noorweghen*), running via Maastricht in Limburg is mentioned. In the charters the name *denemarken* is found three times, twice as a surname and once in a feudal register of count Florens V of Holland, where a man gives his land situated *jn denemarken* to the count and gets it back from him in feudal benefice (*Corpus I*, page 496, line 12; Holland 1285). It is not probable that here a property in Denmark is meant, nonetheless because the remainder of the estates is situated in the diocese of Utrecht.

Further research has solved this problem. Since the 12th century, much land in the diocese of Utrecht had been cultivated on instigation of the bishop. It appears that a number of these cultivations have been named after foreign localities. In the neighbourhood of the cultivation *denemarken*, nowadays called Demmerik, cultivations have been found named after Britain, Bulgaria, Spain and even Cockaigne. It remains vague, however, how this trend to name cultivations after foreign localities came into being (Rentenaar (1984), pp. 81, 108).

Another question, which - at any rate until now - remains unanswered, is whether *Willem van denemarken* (*Corpus I*, page 181, line 5), mentioned in 1270 in Ghent, and *Simon van denemaerke* (*Corpus I*, page 2882, line 16), found in Bruges by the end of the 13th century, came from Denmark or from Demmerik. The second possibility, however, seems the more probable, because Demmerik is much nearer to Ghent and Bruges than Denmark.

The word form *denemarken* is split into two articles for the *VMNW*. The first article (*denemarken (I)*) treats the country name Denmark with reference to the legend of Saint Servatius. The second one (*denemarken (II)*) presents the place name Demmerik, named after the country Denmark, and, under a nuance in meaning, the use of the place name as a surname. The choice to mention the surnames under *denemarken (II)* is explained at the end of the article.

The entries *denemarken (I)* and *denemarken (II)* in the *VMNW*:

#### DENEMARKEN (I)

znw.

#### Denemarken

[1]

0/0/1 *Limburg 1200*. *denemarken*, *ds*.

I 1. **Denemarken, land in Noord-Europa.** || [ende] also gelike. uan sassen te francrike. Ende bit scheppe di des plegen. te denemarken ende te norwegen. *Corpus II (Servas)* p.292, r.12-15 (*Limburg 1200*).

## DENEMARKEN (II)

noun

## Demmerik

[3]

0/3/0 *Gent 1270*. denemaerke, denemarken, *ds. Afgeleid van de landnaam Denemarken*.

**I 1. Demmerik, plaats in de huidige provincie Utrecht.** || gherard van der .a. anderhalve hoeve jn denemarken. *Corpus I p.496, r.12 (Holland 1285)*.

**1.1. Ook in de verbinding van denemarken. Uit Demmerik afkomstig, toenaam van verschillende personen.** || *Willem van denemarken Corpus I p.181, r.05 (Gent 1270)*. Simon van denemaerke *Corpus I p.2881, r.16 (Brugge 1291-1300)*.

*Op grond van de vorm van de toenaam kan niet worden uitgemaakt of de bovengenoemde personen hun naam aan de plaats Demmerik dan wel aan het land Denemarken ontleend hebben. Vanwege het feit dat Denemarken veel verder van Brugge en Gent ligt dan Demmerik, is de verbinding onder dit laatste lemma opgenomen.*

Many places came into being and have been given a name (much) earlier than the 13th century. The description of these place names in the *VMNW* aims more at identification than at disentanglement of the original word meaning. There are, however, places originating from the 12th and 13th century, of which the original meaning of the names is, of course, of outstanding interest for the dictionary. After all, it is obvious that the words used to name places must have had some meaning to the people that used them. This may be shown by the next example.

In the 12th and 13th century a lot of land in Utrecht and Holland was brought under cultivation. In this newly cultivated land various places came into being and got names. Apart from such places as Demmerik, which got named after a foreign location, there are also a lot of places with a name ending in *-wijn*. Blok (1957), p. 23 even speaks of a *wijn*-trend. In Utrecht we find, for example, place names as *hanewijn*, *honswijn* and *scalwijn*, in Holland place names as *blesewijn*, *poedelwijn*, *reewijn*, *rijtwijn* and *schelwijn*. With regard to the origin of the word *wijn* opinions vary. The most widely accepted theory is that Middle Dutch *wijn* in most cases is borrowed from Latin *vicus* with a general meaning 'settlement' and spread through the trade up and down the rivers from the south all over the Low Countries (Blok/Koch (1964), p. 50). This is also the etymology of the word in the *MNW*. Another theory is that the word came to the Low Countries with the Vikings and may thus be of Scandinavian origin (e.g. Fockema Andrea (1954), pp. 81-87). Scandinavian influences on Dutch place naming have, however, convincingly been rejected by Schönfeld (1959) and Blok (1978) and, with regard to Flanders, by Roelandts (1989). Schütte (1976) suggests as origin Germanic *wik*, originally meaning 'fence', which in the course of time develops to 'settlement (of any size)'. In origin it is the same word as the fourth possibility, Germanic *wik* 'bay' (from which Scandinavian *vik* also originates), which gets a special meaning 'wide ditch, canal' in the cultivations in Groningen and Drente. Significant distinction between the cultivations in Holland and Utrecht and those in the north (which are of a later date) is that in the north *wijn* indeed has been recorded as a water name, while in Holland and Utrecht no sign of that has been found.

The possibility that *wijc* in the place names in the 12th and 13th-century cultivations is borrowed from Latin *vicus* 'settlement' is the most common theory and looks very probable at first sight. But there appears to be one problem: two of the place names ending in *-wijc* in these cultivations are recorded preceded by a female article, namely *die schelwijc* (accusative singular; *Corpus I*, page 508, line 43; Holland, the count's chancellery 1280-1287) and *van der poedelwike* (dative singular; *Corpus I*, page 2442, line 29; Koningsveld 1297). A second look makes clear that the others, according to their flexion, can be male as well as female. Latin *vicus* is a male word and the dictionaries give for Middle Dutch *wijc* 'settlement' a male gender too, so the above-mentioned place names may have had another origin. The Germanic words *wik* have a female gender, and especially the one with the meaning 'settlement (of any size)' seems to the point for the cultivations, which can vary in size from one farm to a complete village. However, Germanic *wik* 'bay' might also be a possible origin. Under the influence of Latin *vicus* the meaning of this word may have developed from 'bay' via 'settlement near a bay' to 'settlement (in general)'.

For the 12th and 13th-century place names ending in *-wijc* in the cultivations in Holland and Utrecht the origin thus could very well be Germanic *wik* meaning 'settlement'. This, however, does not mean that all place names ending in *-wijc* share this origin. Some of them are evidently borrowed from Latin *vicus*, as is the case for *wijc* near Maastricht and *wijc* near Tongeren, both in Limburg. The Middle Dutch word *wijc* therefore is a clear example of the fact that (the etymology of) place names can be very interesting for a dictionary. Should the *VMNW* not include names, the existence of a female word *wijc* (< Germanic *wik*) and that of the much more known male word *wijc* (< Latin *vicus*) would have remained unmentioned, because neither of them has been recorded outside onomastic contexts in 13th-century texts.

## Conclusion

On basis of this small selection of examples from the *VMNW* I hope to have shown how valuable onomastic material can be in the description of an earlier language period. Many words have only been recorded as names. The nouns *vlaminic* and *hollander* appeared not only as surnames, but also as first names; to the noun *crudepenninc* a meaning can be added, with the help of other nouns ending in *-penninc*, although the exact relation between the word and the use as a surname remains vague. Research and explanation of place names can clarify earlier texts. The naming of the place Demmerik after the country Denmark is an example of this phenomenon. The etymology of place name elements such as *-wijc* undoubtedly deserves a place in a historical dictionary.

In my opinion a dictionary, based on historical principles, needs to include onomastic material, but I also think it important to distinguish between names and other words, because names have a special function of identification in language, which may cause (slight) differences. By describing a language period in a dictionary on basis of *all* the recorded material - though for practical reasons it may be necessary to restrict the scientific work to a balanced selection - I think linguistic research will gain very much.

## Bibliography

- Blok, D.P. (1957). Naamsveranderingen en modeverschijnselen in de middeleeuwen in Utrecht en Holland benoorden de Lek. In: *Mededelingen van de Vereniging voor Naamkunde* 33, pp. 17-26.
- Blok, D.P. (1977). Toponymische verkenningen in het Land van Woerden. In: *K.N.A.G. Geografisch tijdschrift* N.R. XI, pp. 399-406.
- Blok, D.P. (1978). De Vikingen in Friesland. In: *Naamkunde* 10, pp. 25-47.
- Blok, D.P. en Koch, A.C.F. (1964). De naam Wijk bij Duurstede in verband met de ligging der stad. In: *Mededelingen van de Vereniging voor Naamkunde* 40, pp. 38-51.
- Corpus I. Corpus van Middelnederlandse teksten (tot en met het jaar 1300); Reeks I: ambtelijke bescheiden.* 1977. Uitgegeven door M. Gysseling. 's-Gravenhage. 9 dln.
- Corpus II. Corpus van Middelnederlandse teksten (tot en met het jaar 1300); Reeks II: literaire bescheiden.* 1980-1987. 's-Gravenhage-Leiden. 6 dln.
- Diatessaron Leodiense.* (1970). Edidit C.C. de Bruin. Addita est interpretatio anglica quam curavit A.J. Barnouw. Leiden.
- Falk, H.S. und Torp, A. (1910-1911). *Norwegisch - Dänisches etymologisches Wörterbuch.* Heidelberg. 2 dln.
- Fockema Andrea, S.J. (1954). Vikinger-wijken in Nederland. In: *Mededelingen van de Vereniging voor Naamkunde* 30, pp. 81-87.
- Het Glossarium Bernense. Een vroegmiddelnederlandse tweetalige Latijns-Limburgse woordenlijst.* (1977). Uitgegeven door L. de Man en P.G.J. van Sterkenburg. 's-Gravenhage.
- Het oudste goederenregister van Oudenbiezen (1280-1344).* (1965). (Uitgegeven) door J. Buntinx en M. Gysseling. Dl. 1, tekst. Tongeren.
- Rentenaar, Robert. (1984). *Vernoemingsnamen. Een onderzoek naar de rol van de vernoeming in de nederlandse toponymie.* Amsterdam.
- Roelandts, Karel. (1989). Nordische Einflüsse auf die altflämische Namengebung?. In: *Studia Onomastica, Festschrift till Thorsten Andersson.* Lund, pp. 303-311.
- Schaar, J. van der. (1984). *Woordenboek van voornamen.* Utrecht/Antwerpen.
- Schönfeld, M. (1959). Directe en indirecte wikingervloeden op onze naamgeving. In: *Nagelaten opstellen, Studies op naamkundig gebied.* Amsterdam, pp. 11-44.
- Schoonheim, Tanneke. (1990). Autochthonous names in the *Dictionary of Early Middle Dutch.* In: *Proceedings of the XVIIth International Congress of Onomastic Sciences.* Helsinki. 2 vol., pp. 332-339.
- Schütte, L. (1976). *Wik, eine Siedlungsbezeichnung in historischen und sprachlichen Bezüge.* Köln-Wien.
- Verwijs, E. en Verdam, J. (1885-1952). *Middelnederlandsch Woordenboek.* 's-Gravenhage. 11 dln.
- Woordenboek der Nederlandsche Taal.* (1882-...) Bew. door M. de Vries, L.A. te Winkel, e.a. 's-Gravenhage enz. ... dln.

## Om ord og termer i juridisk språk

Det er ikke lett for vanlige mennesker å lese en lovt tekst og virkelig forstå den. Allmennspråklige ordbøker tar derfor med en del faguttrykk og fagterminologi for å hjelpe lekmannen til å forstå juridiske tekster. Artikkelen viser noen eksempler på hvordan en vanlig ordbok lett kan bli feil dersom en tar med fagspråklige termer.

En av de mest kjente og anerkjente juristene her i landet, Johs Andenæs, skriver i sin lærebok *Innføring i rettsstudiet* (Andenæs 1988):

De fleste juridiske framstillinger bruker en rekke ord og uttrykk som ikke er umiddelbart forståelige for ikke-jurister, men som forfatteren forutsetter at leseren kjenner, og som han derfor ikke forklarer nærmere (s 145).

Andenæs anbefaler derfor den ferske jusstudent å anskaffe seg en juridisk ordbok. Han tar selv opp i læreboka si en del juridiske grunnbegreper og definerer dem etter å ha gitt en forholdsvis omfattende utredning om de enkelte ordenes innhold. Og han konkluderer utredningen sin med:

Rent generelt kan det sies at det i tvilstilfelle ikke er så viktig hvilken etikett man bruker, og at man ikke bør treffe slutninger fra den betegnelse man velger, til de rettsregler som gjelder. Framgangsmåten bør være den motsatte: Først prøver man å klarlegge hvilke regler som gjelder, og så velger man den terminologi som man synes alt i alt gir det mest dekkende uttrykk for realiteten (s 146).

Dette uttrykker en overraskende innstilling til språkets innvirkning på forståeligheten av en tekst. Når det gjelder lovspråk, må ordbruken ha en stor betydning for folks rettsoppfatning, både allmennhetens og juristenes. Særlig hvis vi sammenholder det med hva han sier under innføringskapitlet om tolkning av lover:

Utgangspunktet for tolkningen er hva som ligger i teksten etter en naturlig forståelse på grunnlag av alminnelig språkbruk. Loven retter seg til borgerne, den benytter seg av vårt felles språk, og det er ønskelig at den kan forstås av enhver som er interessert i å skaffe seg rede på sin rettsstilling (s 70).

Etter det man får rede på i en innføringsbok i juss, skal altså menigmann kunne lese en lov og bruke sin naturlige språkforståelse for å finne ut av sine rettigheter og plikter etter lovverket. Men slik er det jo ikke, det viser nettopp Andenæs' eget utsagn der han sier at det

i juridisk forfatterskap ikke er viktig hvilken term man velger, men at det er de juridiske reglene som ligger bak, som er avgjørende.

Andenæs gir videre noen gode råd for juristen som skal tolke lovene, jeg nevner her tre:

- 1 Med hensyn til fagspråk og termer (fra andre fag) er det språkbruken i de kretser som loven retter seg til, som må legges til grunn.
- 2 For anakronismer og andre gamle uttrykk er det den opprinnelige betydningen som må legges til grunn.
- 3 For uttrykk som forutsetter en rettslig *vurdering* (en rettslig standard), kan innholdet endre seg når vurderingsnormene i samfunnet skifter.

Kommunikasjonsmessig ville det være riktig og viktig å tilstrebe et samsvar mellom juridisk innhold og allmennspråklig uttrykk. Men situasjonen vi opplever, er at den juridiske terminologien, eller kanskje rettere ordbruken, anvendes så spesielt at man trenger nærmere kjennskap til juridiske teorier og vendinger for virkelig å forstå tekstene. Og i tillegg vet vi at jurister seg imellom er uenige om hvordan en juridisk tekst skal forstås. Mange ganger virker det som om det er nettopp det jussen dreier seg om. Det betyr igjen at terminologien ikke er helt klar eller entydig, selv for fagekspertene. I en del tilfeller trekker juristene inn språkfolk for å få oppklart hvilket begrepsinnhold som ligger i juridiske termer. Det er i seg selv noe overraskende. Det bør vel være juristene selv som best vet hva de legger i en fagterm, de kan ikke først bruke et uttrykk og deretter kontakte språkfolk (gjærne leksikografer) for å få rede på hvilket betydningsomfang uttrykket har. Situasjonen viser i hvert fall den konflikt som det er mellom fagspråklig og allmennspråklig forståelse av lovt tekstene. Jon Bing (Bing 1986:135) tar opp forholdet mellom allmennspråkets og det juridiske språkets betydning, og han peker også på at det er en viss spenning mellom det innhold som kan følge av alminnelig språkbruk og den juridiske betydningen.

Hvis man skal følge Andenæs' oppskrift for juridisk forfatterskap i at man ikke først og fremst skal legge vekt på termbruken, men velge terminologi etter hva man synes er mest dekkende, vil det bli den enkelte forfatters skjønn og dermed til en stor grad tilfeldighetene som råder innen juridisk terminologi. Innen fagterminologien bør det omvendte være tilfelle. Termene i seg selv er nok mindre viktige, men definisjonen av dem må være helt klare og avgrensede. Det betyr ikke at man ikke trenger være nøyaktig med hvilken "etikett" man bruker i et juridisk forfatterskap. En juridisk forfatter bør kunne støtte seg på hevdvunnen og normert terminologi. Ved innføring av nye termer må juristene selv definere termene og sørge for at de passer inn i den resterende delen av termsystemet.

Skillet mellom allmennspråk og fagspråk er uklart og fortsatt lite utforsket. Men skillet mellom allmennspråkets ord og terminologi er ganske klart, og kanskje nyttig å peke på i denne sammenhengen. Wüster (Wüster 1985) trekker opp skillet mellom allmennspråklige uttrykk og termer i noen grunnleggende punkter omtrent slik:

- I terminologien går man ut fra *begrepene* og har som mål å skape en skarp avgrensning mellom dem.
- Termenes betydning omfatter bare dets denotasjon, konnotasjonene faller vekk.
- Terminologi befatter seg bare med benevnelser for begrepene, ikke med morfologi og syntaks.

- Et fags terminologi er produktet av bevisste, normerende handlinger utført av eksperter, den er ikke resultatet av en fri eller naturlig språkutvikling.

Terminologiarbeid betyr at en hele tiden vurderer forskjellige språklige uttrykk opp mot hverandre og bestemmer hva som er best. Synonymer skal ikke forekomme. Målet er allikevel alltid å finne de mest *hensiktsmessige* uttrykkene, mens man i vanlig språknormering ser på *språkriktighet*, dvs samsvar med den virkelige språkbruken i språksamfunnet.

Termene i lovspråk er oppstått på forskjellige måter. Jeg vil nevne noen av de viktigste her:

- 1 Den spesielle og mest merkbare juridiske terminologien stammer for en stor del fra romerretten eller fra kanonisk rett. Dette er juridisk "arvegods" som har lange tradisjoner, og som det er åpenbart må defineres i faglige oppslagsbøker og leksika. Det er også den typen som vi først og fremst finner der. Som eksempel kan nevnes *bonus paterfamilias* som egentlig betyr "god familiefar", men som i dag benyttes som et uttrykk for om en person har handlet forsvarlig.
- 2 Videre oppstår mange termer gjennom rettspraksis, det vil si at det er ordlyden i viktige lover og domsavsigelser som fester seg og utvikler seg til fagterminologi. I juridisk språk finner vi mange anakronismer; ord og uttrykk kan stamme helt fra norrøn tid. De er blitt stående lenge etter normal levetid fordi de er gått over til å bli ren fagterminologi, selv om de kanskje var allmennspråklige uttrykk da de ble brukt første gang. Eksempel på slike alderdommelige uttrykk er *duk og disk* som uttrykk for en persons bopel, eller *kjæremål* av norrønt *kæra*, klage over, og som i dag bare lever i moderne norsk i juridisk terminologi.
- 3 Juridiske forfattere og lærebokforfattere er ansvarlig for en stor del av det juridiske fagspråket. Det skjer ved at de lager definisjoner for de uttrykkene som brukes. Det definerte blir dermed oppfattet som en fagterm. Et eksempel på det har vi i innledningen til lov om folketrygd § 1-1 og 1-2: "Folketrygdens formål er å gi stønad ved sykdom, legemsfeil, svangerskap og nedkomst, arbeidsløyse, alderdom, uførhet, dødsfall og tap av forsørger. Enhver som er bosatt i riket er *trygdet* etter denne lov." Termen *trygdet* har dermed fått en annen betydning enn allmennspråkets, der vi vanligvis oppfatter en trygdet som en som mottar trygd. En spesiell terminologisering har vi ved legaldefinisjoner, definisjoner som tas inn i selve lovene, og som skal legges til grunn for forståelsen av loven. En legaldefinisjon gjelder bare for den loven den står i. Men oftest er det slik at ord eller uttrykk er brukt på en spesiell måte uten at de er gitt noen nærmere definisjon. Innholdet må da trekkes ut av konteksten.

Ved hjelp av noen eksempler vil jeg prøve å vise hvordan juridisk språkbruk kan atskille seg fra allmennspråket på en uheldig måte, slik at lovttekster og annen juridisk litteratur blir vanskelig å forstå for lekfolk. Jeg vil også peke på hvor farlig det kan være å ta inn fagspråklige uttrykk i allmennspråklige ordbøker fordi en lett kan komme til å ta ut ett eller noen uttrykk fra et helt begrepsfelt, og dermed kan definisjonene lett "halte". Men eksemplene viser også at feil like lett oppstår ved at den juridiske terminologien er usystematisk, og at en allmennordbok dermed lett depreciserer eller feiltolker innholdet i termene.



Få av de juridiske termene er såkalte ekte termer, det vil si ord og uttrykk som vi ikke har i språket ellers. De ekte termene vil være lån (særlig fra romerretten) og anakronismer. Mange termer oppstår ved **neosemantisering**. Det skjer som regel enten ved at allmennspråklige uttrykk får et avgrenset faglig innhold, eller ved at gamle juridiske termer får et nytt innhold. Det kan også være at termer fra andre fag benyttes i juridisk sammenheng og på den måten får en egen juridisk definisjon og dermed blir en juridisk term. Det er bare legaldefinisjonene som til en viss grad er laget etter noe jussfaglig system eller for å dekke hele semantiske felt. Men til og med her er det dessverre slik at det samme uttrykket kan være definert forskjellig i forskjellige lover ved legaldefinisjon.

Svært ofte er det en neosemantisering som finner sted ved legaldefinisjoner. Et eksempel på det har vi ved ordet *tjenestemann*, som er en viktig term i flere lover. En tjenestemann er ifølge flere av de større norske ordbøkene en offentlig funksjonær som ikke er embetsmann, f eks i *Bokmålsordboka* (Landrø & Wangensteen 1986:607) defineres ordet slik:

**tjenestemann** offentlig funksjonær som ikke er embetsmann

Men i lov om offentlighet i forvaltningen fra 1967 defineres tjenestemann som en embetsmann eller annen som er ansatt i statens eller kommunens tjeneste. I og med at uttrykket "eller annen" er brukt, må det tolkes slik at de fleste som er ansatt i statens eller kommunens tjeneste, er tjenestemenn. Det samsvarer med definisjonen i Gulbrandsen *Juridisk leksikon* (Gulbrandsen 1977:213):

**tjenestemann** (offentlig t.) person som er ansatt i statens el en kommunes tjeneste og lønnes av det offentlige. Bestemmelser om hvem som skal regnes til t. gis i tvilstilfelle av Kongen ...

Bing (1986:136) nevner at tjenestemann har forskjellig legaldefinisjon i lov om statens tjenestemenn (1983) og straffelovens (1902) § 121 om offentlige tjenestemenns taushetsplikt. Definisjonen i lov om offentlighet i forvaltningen samsvarer med lov om statens tjenestemenn og Gulbrandsens definisjon.

Hvis vi vil gå videre for å finne ut av betydningen i *Bokmålsordboka*, må vi se nærmere på uttrykkene *funksjonær* og *embetsmann*. De defineres i samme ordbok slik:

**funksjonær** tjenestemann, bestillingsmann; fast ansatt person i privat virksomhet (s 181)

**embetsmann** person som innehar et embete (s 123)

**embete** 1 høyere stilling i statstjenesten 2 oppdrag, tjeneste (s 123)

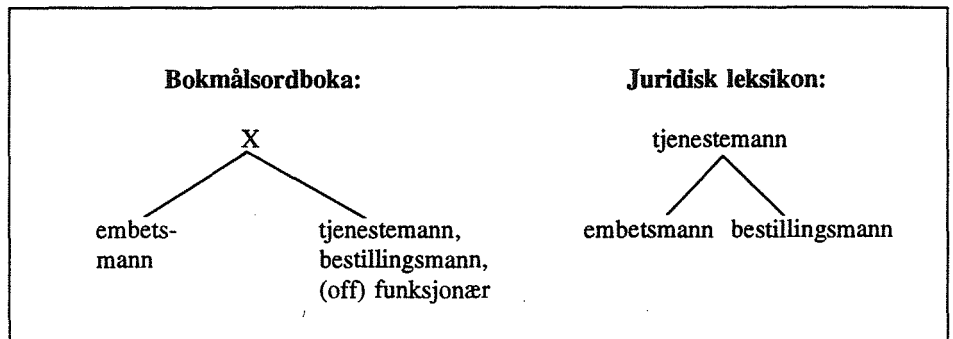
Vi bør ut fra det kunne slutte at en tjenestemann har en underordnet stilling i offentlig virksomhet. Men i juridisk terminologi har vi sett at en tjenestemann er enhver som er lønnet av stat eller kommune. I allmennspråket er det her problematisk å trekke grensen mellom tjenestemann og arbeidstaker. *Store norske ordbok* (Guttu 1991:196) definerer funksjonær slik

**funksjonær** person som har en - ofte forholdsvis underordnet - stilling i offentlig el (større) privat virksomhet og som utfører ikke-manuelt arbeid

Skulle en slik definisjon legges til grunn ved tolkning av loven, ville en måtte stille seg mange spørsmål: Hvem er det som regnes som offentlige tjenestemenn av sjåførere, gartnere, ledere for renholdsarbeidere osv.

Med utgangspunkt i *Bokmålsordboka* ville vi også være usikre på grensen mellom funksjonær og embetsmann, siden en embetsmann har en høyere stilling. Hvor går grensen for en "høyere" stilling i det offentlige? I juridisk terminologi betyr en embetsmann en tjenestemann som er utnevnt av Kongen.

Det viser seg altså at i juridisk terminologi er definisjonen av tjenestemann videre enn i allmennspråket. Dette er illustrert i figur 1.



Figur 1

*Bokmålsordboka* mangler tydeligvis et overbegrep for offentlig tilsatte, mens tjenestemann og bestillingsmann anses som synonyme. I juridisk terminologi er tjenestemann overbegrep, og embetsmann og bestillingsmann er sideordnet med hver sin delbetydning. I praktisk lovanvendelse kan slike skjevheter få uheldige følger for menigmann, da det vil være den juridiske terminologien som gjelder.

Det er sjelden (eller kanskje tilfeldig) at begrepene er dannet etter et semantisk system, slik man gjør det innen terminologi basert på språkvitenskapelige prinsipper, og slik at hele semantiske felt blir dekket. Eugen Wüster (Wüster 1985:5) sier om utviklingen av terminologi:

Die Terminologiearbeit geht nicht von den Begriffen aus, sondern sie untersucht alle Begriffe eines Sachgebietes im Zusammenhang, d.h. alle Teile eines Begriffsystems. Das führt zwangsläufig dazu, dass alle grundlegenden Fachwörterbücher heute "systematisch" angeordnet sind ...

Men det virker som om de juridiske termene oppstår etter hvert som man har bruk for dem. Det kan derfor lett skje at det blir betydningsluker eller betydningsgråsoner, slik som vi ofte finner det i allmennspråket. (Dette er luker eller soner som ikke kan regnes til vaghetsområdet, men kanskje til den betydningsluken som Kit Fine (Fine 1975) opererer med, og som kan liknes med en hvit flekk på kartet, dvs ikke beskrevne innholdssoner.) Slike luker er selvfølgelig uheldige juridisk sett. En juridisk terminologi skapt i Wüsters ånd, foreligger

ennå ikke på norsk, men med utvikling av moderne rettsinformatikk blir behovet stadig større. Wüster sier det slik:

Ein wesentlicher Unterschied zwischen Terminologielehre und Sprachwissenschaft ist darin zu suchen, dass die Terminologielehre Anleihen bei der Logik und bei der Ontologie machen muss und sich mit einer dritten formalen Wissenschaft überschneidet, nämlich der Informationswissenschaft (s 5).

Bare ord og uttrykk fra romerretten er juridisk terminologi som konsekvent kan gjenfinnes i juridiske oppslagsbøker. Termene har selvsagt sitt opphav i latin, og selv om de er oversatt til moderne norsk, er det i oppslagsbøkene henvist til det latinske uttrykket. Juridisk terminologi finner vi forklart i spesialordbøker for juss, og det fins en del juridiske termer i allmennspråklige ordbøker. Jeg har sett litt på hva som er tatt med i *Bokmålsordboka* av juridisk terminologi. Det er jo ofte slik at allmennordbøkene har med faguttrykk og merker av for fagområde med **jur** eller liknende. Dette ordmaterialet er tilfeldig valgt ut, sett fra juridisk synsvinkel. Det er ikke jussfaglige kriterier som gjør at et ord kommer med (for eksempel er ikke ordet *bestillingsmann* tatt med som eget oppslagsord, bare som et synonym til tjenestemann), men at det er blitt forholdsvis frekvent i allmennspråket, eller at ordene på grunn av juridisk neosemantisering blir flertydige. Oppslagsordet får da som regel et eget betydningsnummer for den juridiske definisjonen.

*Bokmålsordboka* har tatt med ordet *forsettlig*, et ord som nærmest bare forekommer i juridisk sammenheng i moderne norsk. Der er koden **jur** gitt, og forklaringen "*f\_ drap* drap som er begått med vilje, men ikke planlagt, t forskj fra *overlagt drap*, *uaktsomt drap*". Under oppslagsordet *overlagt* finner vi forklaringen "tilsiktet, planlagt" og eksemplet "*o\_ drap* t forskjell fra *uaktsomt drap* og *forsettlig drap*". Men at det er en spesiell tolkning av dette ordet i juridisk sammenheng, opplyses ikke, og det spørs om en kan vente at en leser slår opp på *forsettlig* når han er ute etter betydningen av *overlagt* eller *uaktsom*. Men den juridiske spesialbetydningen er like viktig i alle tre termene.

Under oppslagsordet *uaktsom* finner vi definisjonen "uforsiktig, skjødesløs; som skyldes uforsiktighet, skjødesløshet" og eksemplet *u\_t drap*. Det er bare under oppslagsordet *forsettlig* at koden **jur** er lagt inn.

*Juridisk leksikon* definerer slik:

**forsett** er den skyldgrad ved bedømmelse av den subjektive side av en straffbar handling som består i at forbryteren *har tilsiktet* handlingen og dennes følge ... (s 78).

**overlegg** betegner tradisjonelt den sinnstilstand som behersker gjerningsmannen når han er seg bevisst så vel de motiver som kan avholde fra handlingen, som dem som driver til den, og har veiet begge mot hverandre... (s 153).

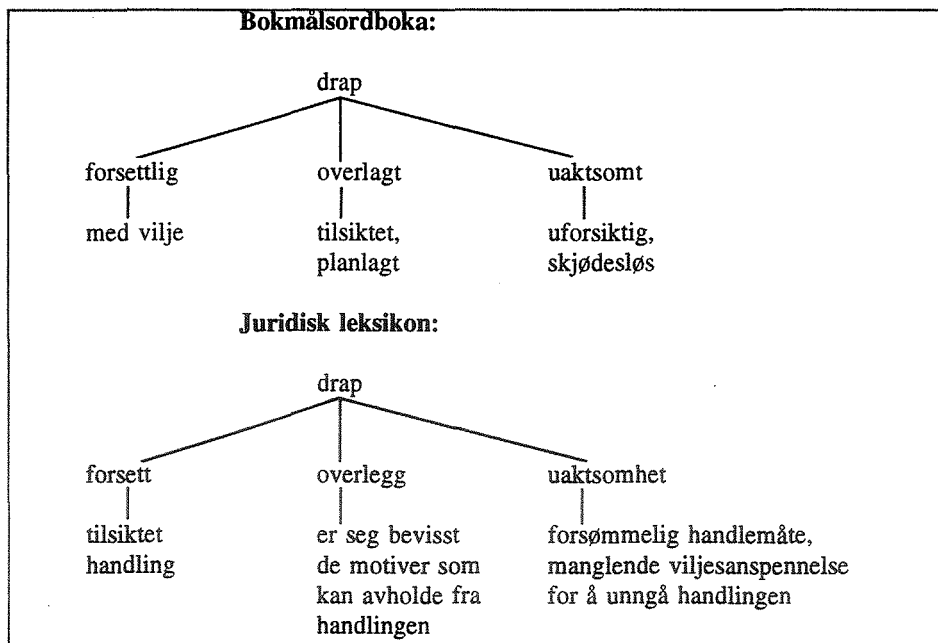
**uaktsomhet** (culpa) uforsiktig el. forsømmelig handlemåte. 1) I strafferett er u. den skyldgrad som i visse tilfelle medfører straff selvom det ikke foreligger **forsett** (s.d.) som er den vanlige subjektive betingelse for ilegging av straff. U. kan sies å foreligge når gjerningsmannen har tilsidesatt den aktpågivenhet han i det enkelte tilfelle må antas å ha hatt plikt til å utvise. Man reagerer mot

den manglende viljesanspennelse; for evneløshet alene blir en mann aldri straffskyldig (s 217).

Denne "manglende viljesanspennelsen" kan vurderes som bevisst uaktsomhet og ubevisst uaktsomhet, men en kan også skille etter grov uaktsomhet og ikke-grov uaktsomhet. Som betydningsnummer 2 står det:

i erstatningsrett er u. utilstrekkelig til å begrunne et erstatningsansvar idet det tas hensyn til den aktsomhet som *gjennomsnittsendividet* i det konkrete tilfelle ville ha utvist, enn til skadevolderens egen utrustning, medmindre denne ikke er abnormt under det vanlige (vanfør, sinnssyk el.l.) (s 218).

Dette er vanskelig å få klarhet i for lekfolk, ikke minst på grunn av språket. Men selv om en ser bort fra den vanskelige språkformen, viser det seg at det betydningsomfanget som ligger i en juridisk term, er forholdsvis vagt beskrevet, det fins ingen klare definisjoner på hva som ligger i dem. Dermed blir det opp til rettsapparatets skjønn å tolke uttrykkene. Rent språklig er det vanskelig å si om det er noen forskjell på å ha tilsiktet en handling og det å ha utført den mens en var seg seg bevisst de motiver som kunne avholde fra handlingen, eller om det forelå en manglende viljesanspennelse for å unngå konsekvensene av handlingen. Stiller man definisjonene i *Bokmålsordboka* og *Juridisk leksikon* opp mot hverandre (figur 2), ser vi at det er et uheldig manglende samsvar mellom den allmennspråklige tolkningen og den fagspråklige.



Figur 2

Det er først når man ser uttrykkene i sammenheng, at det blir så klart for oss lekfolk hvor uskarpe grensene mellom de forskjellige uttrykkene er. Men de juridiske konsekvensene er svært forskjellige. Det er også interessant å merke seg at betydningsskillet er mye klarere i allmennordbøkene enn i fagordbøkene på den måten at termene der dekker hele betydningfeltet fra planlagt drap til uforskyldt drap og blir definert i forhold til hverandre. Det er ikke tilfelle i fagordboka.

Uttrykkene uaktsom - forsettlig - overlagt går igjen i mange sammenhenger i lovspråket, og fungerer som et graderingssystem bl a i forhold til straffeloven. De hører til det som Sundby (1974:200) kaller for *avveiningsmarkører*, dvs ord eller uttrykk som gir anvisning på at en rettslig vurdering skal finne sted når graden av lovbrudd skal avgjøres. På linje med at en innen medisinen snakker om forbrenning av 1., 2. og 3. grad, kunne en også i jussen brukt et liknende termsett, f eks drap av 1., 2. eller 3. grad. Det ville da ha vært lettere å forstå at det her foreligger et spesielt fagspråk, og en unngår forvirring i forhold til allmennspråklig forståelse av uttrykkene. Noe helt annet og språklig uheldig i denne sammenheng er betegnelsen *uaktsom* brukt om drap, parallelt med overlagt og forsettlig. Relasjonen mellom leddene er forskjellig, og et uaktsomt drap krever egentlig et handlende overledd, dvs at vi nå får en aktivisering av selve drapet, mens det som menes selvfølgelig er at det er den som dreper som er uaktsom.

Noe av det vanskeligste og kanskje viktigste i lovspråket er graderingssystemet, og vi kan si at hele lovverket er konstruert over et graderingssystem fra *riktig* til *galt*.

Andenæs drøfter uttrykket *god tro* i motsetning til *slett tro*. For det første er uttrykket *slett* brukt med denne betydning en anakronisme som både har sin egen stilverdi, og som kan skape forvirring særlig hos yngre lesere, men det er en annen sak som jeg ikke vil komme

inn på her. Men det er altså slik at *tro* kan være enten *god* eller *slett*. For eksempel i lov om godtroendeerverv § 1.2 defineres uttrykket *god tro*. Det er han når en mottaker ikke forstår at avhenderen mangler rett og heller ikke burde ha forstått det om han hadde vært så aktsom som det må kunne kreves etter forholdene ved avhendingen, tingens art og omstendighetene ellers.

Man taler altså i loven om aktsom god tro, og dermed skulle man kunne tenke seg at det motsatte er *uaktsom god tro*, men det er ikke noe juridisk begrep. Det samme gjelder under definisjonen av *slett tro*, slik blir det altså ikke symmetri i terminologien. Kan hende fins det ingen tilstand som kan kalles *uaktsom god tro*, men jeg tror det ville vært en fordel innen juridisk metode å ha en slik logisk utbygd graderingsterminologi.

Det har innen jussen nærmest vært en ubestridelig sannhet at lovspråk følger sine egne regler, og at man ikke kan legge allmennspråklige regler og definisjoner til grunn for å tolke lovene. Det spørres om denne sannheten kommer til å bli stående særlig lenge. Peter Goodrich (Goodrich 1987) har gjort en analyse av lovspråkets funksjon i samfunnet, og han konkluderer med at det å se på språk og lover som to atskilte disipliner, er en myte. Røsæg (1992) peker også på dette i sin avhandling *Garantier eller fattigmanns trøst?*:

Språkregler er nødvendige for at den samlingen av ord et utsagn består av, overhodet skal danne ett eller flere meningsfylte utsagn. Så langt vil alle som driver rettslig tolkning bruke språkreglene. Fremstillingen nedenfor går kanskje litt lenger i bruken av dem. Dels skyldes dette et ønske om å nyttiggjøre seg resultater fra språklige drøftelser, og om å se hvor langt problemstillingene i slike drøftelser er parallelle til de juridiske. Men diskusjonen av språkregler er

også forårsaket av et verdistandpunkt om at jussen i størst mulig grad bør legge de alminnelige språkbruksregler til grunn. Det letter trolig forståelsen om man i størst mulig grad kan unngå at det er forskjellige språkbruksregler som brukes i henholdsvis juridiske og andre sammenhenger (s 143).

Mange av de lovene som gis i et moderne demokratisk samfunn, har andre funksjoner enn tidligere lover. Allmennheten er bedre opplyst enn tidligere, og folk forventer at de skal kunne ta teksten i en lov for det de forstår den som. Lovene har forskjellig mål og funksjon og skal fungere i forskjellige kommunikasjonsammenhenger. Mange nyere lover som f eks røykeloven av 1990, er rent holdningsskapende lover. Slik lover vil ha liten eller ingen funksjon dersom de ikke kan leses og forstås av vanlige mennesker.

Allmennspråklige ordbøker bør være forsiktige med å definere fagspråklige termer. Dersom man skal ta med fagspråk, bør det være fordi den fagspråklige termen har gått inn i allmennspråket, eller fordi neosemantisering innen et fagspråk fører til flertydighet av et oppslagsord. Å ta med fagord for fagets egen skyld bør en bare gjøre dersom en har en fullstendig utarbeidet plan der en har laget seg prinsipper for hvilke termer som skal med og hvordan de skal defineres. Ellers kan allmennordboka like gjerne forvirre som oppklare lekfolk om det aktuelle fagområdet. Et fag som juss er svært utsatt i så henseende.

### Litteratur

- Andenæs, Johs. 1988. *Innføring i rettsstudiet*. Oslo  
 Bing, Jon. 1986. Om tolkning av enkeltord. I: *Festskrift til Thorstein Eckhoff*. Oslo  
 Fine, Kit. 1975. *Vagueness, truth and logic*. Synthese 30  
 Goodrich, Peter. 1987. *On legal discourse*. London  
 Gulbrandsen, Egil. 1977. *Juridisk leksikon*. Oslo  
 Guttu, Tor. 1991. *Store norske ordbok*. Oslo  
 Landrø, Marit & Wangensteen, Boye (red). 1986. *Bokmålsordboka*. Oslo  
 Røsæg, Erik. 1992. *Garantier eller fattigmanns trøst? Støtteerklæringer i selskapsforhold av typen "comfort letters"*. Oslo  
 Sundby, Nils Kristian. 1974. *Om normer*. Oslo  
 Wüster, Eugen. 1985. *Einführung in die Allgemeine Terminologielehre und Terminologische Lexikographie*. Kopenhagen

Petter Henriksen

## Forholdet mellom encyklopedologi og leksikografi

Leksikografien og encyklopedologien har hver sine unike arbeidsområder, men en rekke fellestrekk. Artikkelen sammenfatter disse og søker å plassere de to disiplinene innenfor en felles overordnet disiplin - læren om oppslagsverk. Det drøftes om denne disiplinen bør betegnes *leksikografi* i utvidet forstand, eller gis en annen og mer nøytral betegnelse. Behovet for kompetanseutvikling og samarbeid innen feltet øker akutt med det elektroniske oppslagsverks inntog.

Miljøene omkring encyklopedologi og leksikografi lever langt på vei sine egne liv i forhold til hverandre, med beskjeden faglig kontakt. Man kunne fristes til å tro at vi ikke hadde noe å si hverandre.

Riktignok kjenner vi til at våre disipliner gled over i hverandre iallfall fram til forrige århundreskifte. Den første utgaven av *Encyclopædia Britannica*, utgitt 1768-71, hadde til undertittel *A Dictionary of Arts and Sciences, compiled upon a New Plan*. Og da Ephraim Chambers i 1728 utgav sitt hovedsakelig encyklopediske verk *Cyclopaedia, or Universal Dictionary of Arts and Sciences*, er det all grunn til å tro at han selv regnet encyklopedien som et underordnet begrep i forhold til dictionaryen eller ordboka (McArthur 1986: 103). Den dag i dag er det særlig i det franske språkområdet tradisjon for å innlemme encyklopedien i leksikografiens virkefelt.

Det viser seg i dag at våre miljøer arbeider parallelt med en rekke felles oppgaver, for eksempel registrering av nyord, innhenting av opplysninger om uttale og etymologi, og framskaffelse av tekniske hjelpemidler og systemer for redaksjonelt arbeid. Dessuten møtes vi av store felles utfordringer i forbindelse med elektronisk publisering av oppslagsverk. Det kunne se ut som om de elektroniske mediers inntog vil reversere noe av våre miljøers drift fra hverandre. Dette vil jeg komme nærmere inn på nedenfor.

Forhold som disse taler for en drøfting av forholdet mellom våre disipliner. Vi må undersøke om det er fruktbart å etablere en *felles faglig overbygning* over disiplinene, og om det er hensiktsmessig å etablere *fora for samarbeid*.

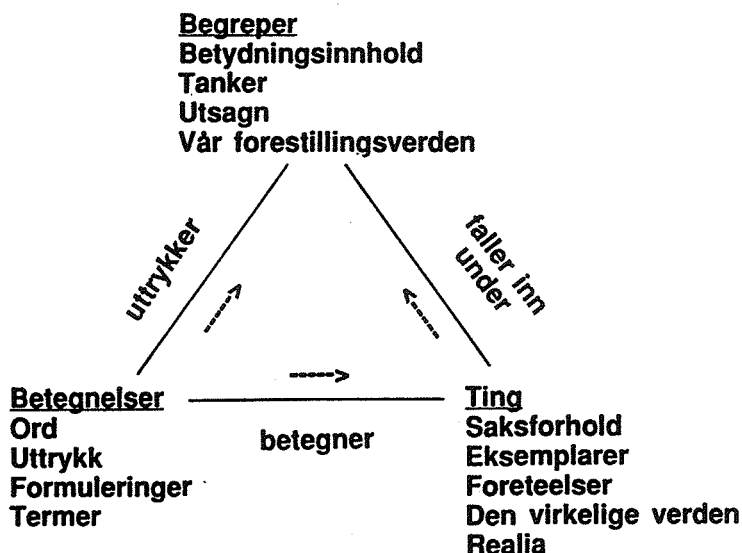
### Slektskap

Ifølge hevdvunne forestillinger beskjeftiger både leksikografien og encyklopedologien seg med *alfabetiske oppslagsverk*. I ordbøker er emnet (gjenstanden for artikkelskriving) *ord* (språklige enheter), i encyklopedier ting og foreteelser, *verdens tilskikkelser*, realopplysninger.

Begrepslærens *semiologiske trekant* (figur 1) skiller mellom (a) betegnelser, (b) begreper og (c) ting eller eksemplarer, som plasseres i trekantens respektive hjørner. Vi ser at ordbøkernes gjenstand befinner seg i nedre venstre hjørne, encyklopedienes i nedre høyre.

Med støtte i dette kan vi gjerne si at encyklopedologien og leksikografien er *søsterdisipliner*, med et klart atskilt utgangspunkt.

### DEN SEMIOLOGISKE TREKANT (OGDENS TREKANT)



Figur 1

Ser vi nærmere på saken, er det ikke fullt så enkelt. Ordbøker inneholder opplysninger ikke bare om betegnelser. De inneholder opplysninger også om *ting* (nederst til høyre i den semiolegiske trekanten) - mange av oppslagsordene og ordforklaringene dreier seg om realopplysninger. Ett av utallige eksempler er artikkelen i *Bokmålsordboka* (Landrø/Wangenstein 1986) om jomsviking:

*jomsviking* (norr *jómsvíkingr*): viking på den danske vikingborgen Jómsborg, trolig på øya Wollin ved munningen av Oder / *ennå er ikke alle j-er døde* (etter Snorre) *ennå er det håp*.

Dette er langt på vei en rent encyklopedisk artikkel, selv om selvfølgelig *formålet* med artikkelen er leksikografisk, nemlig å forklare uttrykket "ennå er ikke alle jomsvikinger døde".

I det engelske språkområdet finnes en rekke verk med betegnelsen "dictionary" som byr på hyppige innslag av encyklopedologisk art. I *The American College Encyclopedic Dictionary* (Barnhart 1952) står for eksempel følgende artikkel om Italia:

*Italy* (it'æli), n. a republic in S Europe, comprising of peninsula S of the Alps, and the islands of Sicily, Sardinia, Elba, etc.: a kingdom, 1870-1946. 45,527,000 pop. (est. 1942); 119,772 sq.mi. (1945). Cap.: Rome. Italina, Italia.



Et mer lokalt eksempel finner vi i *Tanums Store Rettskrivningsordbok* (Sverdrup m/fl. 1983):

4122 *Fiskå* (poststed, Strand kommune, Rogaland).

Dette må sies å være uhemmet encyklopedologi.

Dessuten regnes leksikografien å omfatte slike ordbøker som inndeler universet ikke på betegnelsnivå, men på begrepsnivå, altså tredje hjørne i den semiologiske trekanten. Det gjelder terminologiske ordbøker (for eksempel *Norsk Dataordbok*, Hofstad m/fl. 1987), tesauruser og synonymordbøker. Dette er slike ordbøker som i *Handbok för lexikografi* kalles de *onomasiologiske* (Svensén 1987: 17).

Dermed kan ordbøker klassifiseres slik:

- Ordbøker med *betegnelser* som emne
  - Med appellativer som emne
    - Med rene språklige definisjoner
    - Med henvisning til realia
  - Også med *proprier* som emne
- Ordbøker med *begreper* som emne: tesauruser, synonymordbøker, terminologiske ordbøker o.l.

Encyklopedier på sin side *skal* inneholde artikler fra alle tre hjørner i trekanten:

- Betegnelser/ord, f.eks. fagord, fremmedord, "vanskelige" ord
- Begreper
- Ting/eksemplarer, f.eks. personer, steder, begivenheter

Når *Lademanns Leksikon* (Langer m/fl. 1986) kun har følgende forklaring under oppslagsordet *præsenteren*, er det *begrepet* og ikke betegnelsen presentering som er emnet. For øvrig ser etologien her ut til å ha fått bre seg på bekostning av andre områder:

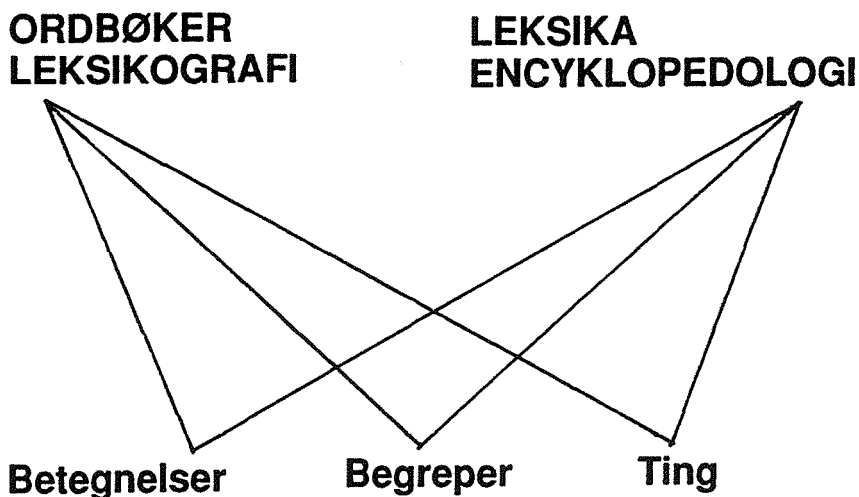
*præsenteren*: hunnabers fremvisning av deres bagdel; et led i seksualadfærden. Det har vist sig, at alle aber uanset køn og alder indtager sammen kroppsstilling over for en sterkere part, hvorfor adfærden har utviklet sig til at være et led i abers sociale rangorden.

Mange av leksikonenes fremmedordartikler er på sin side rene ordboksartikler på betegnelsnivå, f.eks. (fra *Damms Store Leksikon*, Henriksen m/fl. 1985-90):

*intervenere* (av latin, egentlig 'komme imellom'): gripe inn.

Og rett som det er støter vi på artikler i leksikon som beveger seg ubesværet mellom alle tre hjørner i trekanten. Under oppslagsordet *musikk* i *Aschehoug og Gyldendals Store Norske Leksikon* (Tveterås m/fl. 1978-81) finner vi behandling av musikk som betegnelse (uttale, etymologi, definisjon), som begrep ("opprinnelig samlebetegnelse på de kunster (eller kyndigheter) som kommer i stand under Musenes medvirkning ..."), og som ting, fenomen - tonekunst - med en historikk over den europeiske kunstmusikk o.l.

I praksis har altså så vel ordbøker som encyklopedier rikelig med innslag fra alle tre semiologiske nivåer, og det som kunne virke som et ukomplisert søskenforhold mellom våre disipliner, fortøner seg mer sammenfiltret (figur 2).



Figur 2

### Felles overbygning

På bakgrunn av dette kan det være forsøket verdt å se de to disiplinene *under ett*. Er det nok fellestrekk dem imellom - og nok fellesinteresser mellom de respektive miljøer - til å forsvare en grad av faglig sammenknytning? Det er her - vel å merke - ikke snakk om om noen sammensmeltning av fagene. Begge har unike virkefelt som den andre mangler forutsetning til å ta del i. Et sentralt eksempel er leksikografiens tilknytning til lingvistikken, som alltid vil være et mørkt fastland for encyklopedologen. Like fullt må vi undersøke muligheten for en *felles faglig overbygning*.

Det begge miljøer driver med, kan kalles for *oppslagsverk*. Kjernen i definisjonen av oppslagsverk i *Bokmålsordboka* er denne:

*oppslagsverk*: bokverk med alfabetisk ordnede artikler om ulike emner.

Dette er en grei og folkelig definisjon, som imidlertid bør suppleres noe for å være saklig dekkende. Definisjonens genus proximum - bokverk - må nok sies å være i snevreste laget, dermed utelukkes nemlig alle elektroniske produkter, dessuten ikke-boklig informasjon som lyd og video. Trolig er det bedre å innføre som genus proximum *informasjon*.

På samme bakgrunn kan også begrepskjennetegnet *med artikler* virke noe snevert. Audiovisuell informasjon bryter klart med vår tradisjonelle forestilling om en artikkel.

definisjonen inneholde ytterligere et begrepskjennetegn, som går ut på at artiklene har *stikkord* (kjennetegn, identifikatorer, "knagger") som gjør dem gjenfinnelige.

Et forslag til en *presis* definisjon blir dermed dette:

Et oppslagsverk er informasjon som er strukturert etter emne i atskilte enheter med én eller flere identifikatorer, og ordnet med sikte på gjenfinning.

Et forslag til en *kortfattet* definisjon blir dette:

Et oppslagsverk er en samling stikkordbåret informasjon, strukturert med sikte på gjenfinning.

Oppslagsverk kan *deles inn* etter ulike kriterier, for eksempel disse:

- Etter *ordningsmåte (aksessmåte)*: alfabetisk, tematisk (faglig), kronologisk, geografisk, alfabetisk etter ordendelse m.fl.
- Etter *semiologisk nivå* for artiklenes emne: betegnelse, begrep, ting/eksemplar.
- Etter *bredde/stikkordutvalg* (også kalt "scope"): f.eks. bare språkets allmenne ordforråd, også fremmedord, fagord.
- Etter *dybde/opplysningstyper*: f.eks. definisjon, uttrykk, typologi, oversettelse, historikk, etymologi, synonymer, bilder, lyd.
- Etter *normativitet/formål*: deskriptivt, normativt.

Ved noen kombinasjoner av disse egenskapene får vi det vi kaller en typisk ordbok, ved andre det vi kaller en typisk encyklopedi. Ved noen kombinasjoner får vi andre typer oppslagsverk igjen, som kanskje vil passe til betegnelser som *atlas* eller *historiebok*, eller som vi kanskje ikke har treffende betegnelser for i det hele tatt.

Slik kan vi som leksikografer og encyklopedologer posisjonere oss innenfor et felles fag - læren om oppslagsverk.

### Felles interesser

Hvilken interesse skulle våre to miljøer ha av å se på våre felt samlet, som én disiplin, og av å samarbeide?

*Styrking av tradisjonell kompetanse*: Fagfolkene i encyklopedi-bransjen driver til daglig med ordboksarbeid og skriver en lang rekke leksikografiske artikler, men har begrenset kompetanse innen feltet. De trenger økt kontakt med den etablerte leksikografien for å bli dyktigere innen dette området.

Trolig kunne også leksikografer ha noe å lære av encyklopedologer, til hjelp når man beveger seg inn i gråsonen mot encyklopedologi.

*Miljømessig supplering*: Leksikograf-miljøet er tradisjonelt et mer akademisk miljø enn encyklopedolog-miljøet, og mindre kommersielt. Mange encyklopedologer føler de trenger tilførsel av leksikografiens *ånd* - dens akademiske, vitenskapelige tilnærming til sitt felt, med teoriopbygging, lærebøker, utdanning og konferanser. På den annen side vil muligens leksikografien ha godt av ytterligere påvirkning fra leksikonforlagenes kommersielle miljø, fra næringslivet, ikke minst dets kultur for administrasjon og realisering av prosjekter.

*Utvikling av ny kompetanse, særlig i forbindelse med elektronisk publisering:* Den institusjonen som solgte flest ordbøker i 1990, skal ha vært et firma som i utgangspunktet er teknologisk - Franklin, og det med produkter som ikke er bøker, men lommedatamaskiner. Elektroniske oppslagsverk er avgjort i frammarsj.

Kritiske suksessfaktorer i det elektroniske oppslagsverk er:

- Informasjonsarkitektur
- Brukergrensesnitt
- Multimediabehandling

Dette er helt andre suksessfaktorer enn vi har hatt i de tradisjonelle bokprodukter, og som leksikografi- og encyklopedologi-miljøene har spesialisert seg på fra gammelt av. Disse nye suksessfaktorene er *felles* for våre to disipliner - sett fra teknisk systemsynsvinkel er ordbøker og encyklopedier nokså like. Når forleggere sitter samlet og diskuterer emnet *SGML og oppslagsverk*, diskuterer man underforstått ordbøker og leksika i fellesskap.

Her er altså nye kompetanseområder som de to miljøene må samarbeide om, utvikle nye teorier for, og utvikle nye teknikker for. Det er svært viktig at det er vi som redigerer oppslagsverkene, som er toneangivende i dette arbeidet, og ikke de som leverer de tekniske og programvaremessige løsningene.

*Informasjonsarkitektur:* Med hypertekst og fri, mangedimensjonal aksess til databaser åpnes helt nye muligheter for å bygge opp oppslagsverk. Vi får ordbøker og leksika med inngang til artiklene ikke bare alfabetisk etter stikkord, men også f.eks. tematisk (alle nyord, alle maritime uttrykk, alle tabuord), etter etymologisk opphavsspråk (alle lånord fra mellom-nedertysk), osv. Ekspertise på dette felt er vesentlig - risikoen er stor for å lage uløselige informasjonsflokker.

*Multimedia:* Våre to miljøer skal konkurrere med leverandører som tilbyr oppslagsverk med både tekst, bilder, lyd og video. I *Bokmålsordboka* (Landrø/Wangensteen 1986) lyder definisjonen av *dompap* slik:

*dompap:* spurvefugl i finkefamilien med kort, tykt nebb, hodehette og stupeflekk, der hannen har sterkt rød underside, *Pyrrhula pyrrhula*.

Det er tvilsomt om fremtidens ordbok ikke også vil gi visuell og auditiv informasjon om dompapen, f.eks. et opptak av dens sang og en videosekvens med dens opptreden i fugleneket. Det er et spennende spørsmål hvilket miljø og hvilken bransje som vil utvikle slike ordbøker.

*Brukergrensesnitt:* Viktige utfordringer knytter seg til utforming av skjermbildet for at brukeren skal kunne orientere seg i så mye uensartet informasjon uten å forvirres. Sentrale spørsmål knyttes til valg av grafikk og ikoner, og hvordan brukeren skal hjelpes på vei så man ikke navigerer seg bort. I fellesskap må vi som lager og kjenner oppslagsverk utarbeide prinsipper for dette.

Elektronisk publisering vil med andre ord trolig påføre oss flere felles disipliner. Det er et åpent spørsmål om den ikke også vil påføre oss flere *felles produkter*. Når framtidens optiske medier tilbyr utgiveren vesentlig større plass enn oppslagsverk i dag - også større plass enn på dagens CD-ROM-plater - vil behovet for seleksjon bli vesentlig mindre. Det blir lite behov for å skille mellom en vanlig ordbok (med få fremmedord) og den særegne ordbokstypen *fremmedordbok* - vi kan få begge deler i ett, langs samme alfabet. Og vi vil - med den økende konkurransen innen elektronisk publisering - måtte tilby publikum såkalte

*bredspektrede oppslagsverk*, alt i ett-produkter, der så vel ordbokas som encyklopediens innhold er med. Vi kan tenke oss - ut fra brukerens behov - en artikkel under oppslagsordet *malaria*, med en slik meny:

- 1 Uttale
- 2 Etymologi
- 3 Bøyning/ordklasse
- 4 Definisjon
- 5 Eksempler
- 6 Symptomer
- 7 Behandling
- 8 Utsikt til helbredelse
- 9 Historikk, utbredelse
- 10 Interesseorganisasjon
- 11 Bibliografi
- 12 Illustrasjon
- 13 Oversettelse (engelsk, tysk, fransk)

Bredspektrede oppslagsverk vil ventelig tvinge seg fram. Informasjonsleverandøren som stiller seg slik til at brukeren må bytte ut CD-ROM-plate når man skal gå fra definisjon til symptomer, vil tape i kampen om det kjøpende publikum.

Rent pragmatiske hensyn tilsier at vi bør se på våre disipliner *mer under ett*, med én felles faglig overbygning.

### Felles betegnelse - referensologi?

Hvilken betegnelse bør vi bruke om paraply-disiplinen leksikografi-encyklopedologi?

Det er prinsipielt tre muligheter: å utvide betydningen til en av de to eksisterende betegnelser leksikografi eller encyklopedologi, eller å finne på en ny betegnelse. En drøftelse:

*Utvidelse av betegnelsen encyklopedologi (Alternativ A):* I mange encyklopedologers øyne er begrepet encyklopedi overordnet begrepet ordbok, idet encyklopedier inneholder så mye ordboksstoff. I pakt med dette ville betegnelsen *encyklopedologi* være den korrekte for paraply-disiplinen. Vi kunne skille mellom *ord-encyklopedologi* (= leksikografi) og encyklopedologi generelt. Betegnelsen er imidlertid tungvint og lite kjent, og det spørs også om leksikografene ville finne seg til rette under den.

*Utvidelse av betegnelsen leksikografi* til å dekke alle typer stikkordbårede oppslagsverk, også leksikon: Dette er det allerede etablert en viss praksis for, ikke minst i engelsk- og fransktalende land. *Bokmålsordboka* (Landrø/Wangensteen 1986) har som annedefinisjon på leksikografi: ordboks- el. leksikonarbeid.

Tom McArthur (en proklamert leksikograf) gjør seg til talsmann for denne løsningen (McArthur 1986: 109).

"From all this, therefore, I would conclude that the term 'lexicography' *can* legitimately be stretched to cover both ends of the continuum."

Han foreslår videre følgende termpar (analogt med makro-økonomi og mikro-økonomi) for studiet av oppslagsverk generelt (alternativ B):

- 1 *micro-lexicography*, which deals with the world of words and the wordbook proper (which in most instances is an alphabetic dictionary)
- 2 *macro-lexicography*, which shades out into the world of things and subjects, and centres on compendia of knowledge (which in most instances are encyclopedias, which in most instances nowadays are also alphabetic)

Disse betegnelsene er også tungvinte. En mer lettvtint måte å tøyse leksikografibegrepet på, er denne (alternativ C):

- Leksikografi som paraply-betegnelse, omfattende:
  - Tradisjonell leksikografi (pleonastisk kalt *ord*-leksikografi der det er behov for presisering)
  - Tradisjonell encyklopedologi (*real*-leksikografi)

*Ny betegnelse:* Fellesnevneren for våre disipliner er oppslagsverk, på engelsk *reference works*, på fransk *ouvrages de référence*. Når vi snakker om læren om oppslagsverk, ligger derfor et ord som *referensologi* en på tungespissen (alternativ D). En ulempe er at betegnelsen er ukjent. På den annen side er den ikke belemret med tradisjonelle forestillinger. Og den dekker nokså presist det begrepet vi søker en betegnelse på. Altså:

- Referensologi som paraply-betegnelse, omfattende:
  - Leksikografi
  - Encyklopedologi

Alternativ E: Paraply-betegnelsen *leksikologi*. Dette er imidlertid problematisk, da ordet har en helt annen betydning i leksikografiske kretser.

### Konklusjon

Som encyklopedolog innordner jeg meg gjerne leksikografien og går inn for alternativ B eller C. Jeg er imidlertid usikker på hva leksikografien selv vil. I løpet av konferansen har jeg støtt på såvel puristiske som pragmatiske holdninger til dette spørsmålet. Jeg håper det leksikografiske miljøet selv kan komme ytterligere til klarhet om dette sentrale spørsmålet: *Bør leksikografien i vid forstand også beskjeftige seg med studiet av den reelle verden?*

I påvente av dette virker det naturlig å ty til betegnelsen *referensologi* for å dekke paraply-disiplinen leksikografi-encyklopedologi.

Uansett betegnelse håper jeg på et fruktbart samarbeid mellom leksikografer og encyklopedologer i årene framover.

**Litteratur**

- Barnhart, Clarence L. 1952.: *The American College Encyclopedic Dictionary*. New York.
- Henriksen, Petter m/fl. (redaksjon). 1985-90: *Damms Store Leksikon*. Oslo.
- Hofstad, Knut - Løland, Ståle - Scott, Per. 1987: *Norsk Dataordbok*. Oslo.
- Langer, Torben W. 1986: *Lademanns leksikon*. København.
- Landrø, Marit Ingebjørg og Wangensteen, Boye (redaksjon). 1986. *Bokmålsordboka*. Oslo.
- McArthur, Tom. 1986. *Worlds of Reference*. Cambridge.
- Svensén, Bo. 1987. *Handbok i lexikografi*. Stockholm.
- Sverdrup, Jacob - Sandvei, Marius - Fossetøl, Bernt. 1983: *Tanums store rettskrivningsordbok - Bokmål*. Oslo.
- Tveterås, Egil m/fl. 1978-81: *Aschehoug og Gyldendals Store Norske Leksikon*. Oslo.





Del 4



Hans Jonsson

## De stora historiska ordböckernas särart

*Svenska Akademiens ordbok* (SAOB) tillhör gruppen *de stora historiska ordböckerna* och har sina närmaste likar i *New English Dictionary*, *Grimm, Deutsches Wörterbuch* och *Woordenboek der Nederlandsche Taal*. Denna typ av ordböcker behandlar mycket utförligt inom samma pärmar samtidspråket och språket bakåt i tiden ned till gränsen mot ett avgränsbart äldre språkskede, i SAOB:s fall fornsvenskan. Denna sambeskrivning av en så lång språkperiods ordförråd ger SAOB och dess likar en klar särart, när man jämför dem med å ena sidan (hand)ordböcker för nuspråket, å andra sidan ordböcker för en avgränsad äldre språkperiod. Klarast framträder nog särarten i de fall där nutida och äldre betydelsestruktur "krockar" hos ett ord.

Den som slår upp ett "vanligt ord" (många gånger även, för den delen, ett ovanligt ord) i *Svenska Akademiens ordbok* - t.ex. *ande*, *av*, *björn*, *hård*, *mur*, *på*, *sitta*, *stam* - utan att vara väl förtrogen med verket, överraskas nog (främst) av artikelns längd och av den detaljerade betydelsebeskrivningen resulterande i ett flertal, många gånger ett otal, moment/undermoment.

Det finns mycket att säga om kvantitativa och graduella olikheter mellan *Svenska Akademiens ordbok* och andra svenska ordböcker. Den här uppsatsen ägnas dock åt den artolikhet som existerar. Men först en kort presentation av medlemmarna i gruppen de stora historiska ordböckerna.

### 1. SAOB, NED, GRIMM, WNT - och ODS

*Svenska Akademiens ordbok* (SAOB) är en stor (ännu icke fullbordad) historisk ordbok över svenska språket från 1521 och framåt.

SAOB har sina närmaste motsvarigheter i: *New English Dictionary* (NED) som behandlar i första hand engelskan från 1150, dvs. medelengelska och nyengelska, men i aktuella fall följer orden ned i fornengelskan; *Grimm, Deutsches Wörterbuch* (GRIMM) som behandlar tyskan från början av 1500-talet, dvs. nyhögtyska; *Woordenboek der Nederlandsche Taal* (WNT) som behandlar nederländskan från början av 1500-talet, dvs. nynederländska.

SAOB, NED, GRIMM och WNT överensstämmer ingalunda helt i redigeringsprinciper och beskrivningssätt eller i fråga om principer för ordurvalet. - Förenklat kan nog likheterna och olikheterna summeras så att GRIMM intar en särställning med sitt resonerande (kåserande) analysätt och sitt långtgående utelärande av främmande ord. Av de tre övriga står NED och SAOB varandra närmast.

Alla fyra företer dock följande likhet konstituerande *den stora historiska ordboken*: De behandlar inom samma pärmar samtidspråket (nuspråket) och språket bakåt i tiden ned till gränsen mot ett (någorlunda) avgränsbart äldrespråk: fornsvenska (dvs. svenskt medeltids-språk) resp. medelhögtyska och fornengelska (dock följs, som nämnts, nyeng. och meng. ord

ned i forneng.) och medelnederländska. - *Ordbog over det danske Sprog* (ODS), som behandlar danskan från 1700 till 1958, kan också räknas hit. Dock avviker den från gruppen SAOB, NED, GRIMM och WNT i det att den inte behandlar ett helt "nyspråk", dvs. inte hela nydanskan (från omkring 1500).

## 2. SAOB

De senaste enspråkiga svenska handordböckerna är *Illustrerad svensk ordbok* (ISO), 3:e revid. uppl. 1964, och *Svensk ordbok* (SO), 1986. Den jämförelse jag här gör mellan ISO och SO å ena sidan och SAOB å den andra är relevant för förhållandet överhuvud mellan den historiska ordboken och den enspråkiga ordboken av handbokstyp. Såväl ISO/SO som SAOB är enspråkiga svenska ordböcker. Men medan ISO/SO beskriver samtidsspråket nusvenska utgör SAOB en lexikalisk sambeskrivning av samtidsspråket nusvenska (mera korrekt av vad som är "nusvenska" när respektive del av ordboken skrivs) och äldre svenska ned till 1521.

2.1 Man kan då börja med konstaterandet att SAOB till skillnad från ISO/SO är historisk. Och den är historisk inte bara som en automatisk följd av beskrivningsuppgiften. Den är det också så att säga aktivt, i praktiskt taget alla avseenden.

En etymologi inplacerar ordet i ett svenskt och/eller nordiskt sammanhang och i regel finns det också såväl germanska som andra indoeuropeiska anknytningar. Etymologin ger en extern historisk ram åt det behandlade ordet. Främst att observera är nog att den i hög grad ger ledning beträffande vilken betydelse som är äldst (och skall bilda moment 1) hos ett polysemt ord liksom ofta överhuvud för bedömningen av (huvuddragen i) betydelseutvecklingen.

All konkurrens på innehållsplanet dateras i SAOB. Det gäller böjnings- och formvariation liksom genusvariationen hos substantiven. Så långt det går placeras momenten, hos ett ord med flera betydelsemoment, i historisk ordning. När förhållandet mellan olika moment inte ger sig (tämligen) självt hjälps i regel läsaren av korta antydanden eller upplysningar (i etymologiska parenteser) inne i artikeln. I aktuella fall upplyser också etymologiska parenteser vid sammansättningsartiklar om förhållandet mellan huvudord och sammansättningsförled ("namngivningsgrunden"). Nämnas kan också att litteraturexemplen som illustrerar definitionerna ges i kronologisk ordning samt att det första belägget på ett moment eller undermoment är det äldsta i ordbokens samlingar.

2.2 Det samtida svenska ordförråd som såväl ISO/SO som SAOB, behandlar saknar inte kronologisk (dvs. historisk) spännvidd. Där finns ord och ordanvändningar som har prägeln "ålderdomligt språk" (och framför allt används av äldre språkbrukare) liksom ord och ordanvändningar som har prägeln "nytt språk" (och framför allt används av yngre språkbrukare).

I SAOB tillkommer emellertid som redan påpekats den artegna, historiska uppgiften att sambeskriva nusvenskans ordförråd med svenskt ordförråd ända ned till början av 1500-talet. Ett snitt genom svenskt ordförråd i slutet av 1800-talet skulle nog blottlägga relativt lite som 1991 års brukare av svenska, i varje fall äldre sådana, är helt obekanta med. Mycket skulle framstå som ålderdomligt men likväl vara hört hos den äldsta generationen språkbrukare eller hos avlidna gamla språkbrukare vars språk man minns, eller vara att finna i äldre ännu levande litteratur. Med andra ord förhåller det sig nog så att det allra mesta av de senaste 100 årens svenska lever åtminstone i någon utsträckning hos dagens språkbrukare.

Om man tänker sig ett antal ännu äldre snitt genom svenskt ordförråd, låt oss säga från första delen av 1800-, 1700-, 1600- och 1500-talen torde genomgående varje sådant snitt innehålla större inslag av stendöda ord och ordanvändningar än det närmast föregående (dock känner jag mig osäker beträffande förhållandet mellan 1600- och 1500-talssnitten). Inslagen av "annat språk" än nuspråket blir m.a.o. (rätt självklart) större och större ju längre bakåt man kommer i beskrivningsperioden.

**2.3** I åtskilliga fall finner man i SAOB så att säga absoluta olikheter på ordplanet mellan nusvenska och äldre svenska. Dvs., man finner särnusvenska ord, t.ex. *raggare*, *service*, *set*, *strul* och säräldresvenska sådana, t.ex. *messertygare* (en gravstickel), *scherwentsel* (en kortspelsterm), *skött* 'snart o.d.', *stelen* 'stel'. I fråga om sådana artiklar kan man föreställa sig SAOB som enbart en alfabetisk sammanställning av en nusvensk och en äldresvensk ordbok, båda av SAOB:S karaktär bortsett från perioduppdelningen.

Mycket vanligare är dock artiklar behandlande ord som nuspråket och äldrespråket har gemensamt. Praktiskt taget alla sådana har i någon utsträckning förändrats till sin användning. Rätt ofta rör dock inte förändringarna betydelsen i strängare mening utan stannar vid skillnader i fraseologi eller i en och samma betydelses nyanser eller specialfall (förutsatt att man kan avgöra vad som är *betydelse*, *nyans*, *specialfall*). Sådana ord utan betydelseförändring i strängare mening är t.ex. *matt*, *matta*, sbst., *människa*, *mjöl*, *mjölk*, *mången*, *stena* (att döma av resp. artiklar i SAOB).

Betydligt oftare har dock en verklig förändring av ordets betydelseinnehåll inträffat. Ibland är därvid förändringen av den karaktären att den inte påverkar den del av ordinnehållet som är gemensamt för nuspråket och äldrespråket. Verbet *segra* hade i äldre tid även betydelseorna 'besegra' och 'triumfera'. Betydelsen 'vinna seger' har knappast förändrat karaktär, fått annan ställning i ordförrådet, andra oppositioner, på grund av att betydelseorna 'besegra' och 'triumfera' försvunnit. På samma sätt påverkas inte det för nutida och äldre svenska gemensamma betydelseinnehållet hos *konung* (*kung*) av att de äldre betydelseorna 'bivise' (se SAOB K 2294) och 'regulus', 'antimon' (metallurgiska termer; se SAOB K 2295) försvunnit. Inte heller har uppkomsten av en ny betydelse 'gaspedal' hos *spik* (se SAOB S 9633) gett någon ny dimension åt 'järnten som slås in i trä o.d.', den betydelse som funnits under hela beskrivningsperioden. Artiklarna för *slank*, *konung*, *spik* kunde tänkas tillkomma genom att en nusvensk ordbok av SAOB:s typ kompletterats moment för moment med språkprov från en äldrespråksordbok av SAOB:s typ, varefter särmoment från äldrespråksordboken införs intakta.

I många fall innebär dock inte artikelskrivandet i SAOB en sådan okomplicerad sammanställning av motsvarande artiklar i en tänkt nusvensk och en äldresvensk ordbok. Ofta är det nämligen så att ordet i nusvenskan och äldresvenskan visserligen kan ha samma referentiella användning men den semantiska strukturen har ändrats. *Slank* t. ex. är i nusvenskan ett adjektiv med den egentliga betydelsen 'om person: (midje)smal och smidig el. dyl.' och en bildlig användning om träd el. dyl., medan den egentliga betydelsen i äldrespråket var 'om person el. träd el. dyl.: smal och smidig el. slänig' (se Jonsson 1983:149 f.). Med person eller djur som subjekt betydde *springa* ursprungligen 'ta el. göra språng el. skutt, hoppa' medan löpa hade *springa*'s nu gängse innebörd. I viss utsträckning används fortfarande *springa* i fråga om språng eller hoppande, t.ex. i *springa åt sidan*, *springa upp* (från sittande eller liggande ställning), eller som förled i sammansättningarna *springbock*, *springhare*. Sådana användningar hänger kvar i nuspråket (som relikter och) under anslutning till den gängse betydelsen hos *springa*.

I fall som *slank* och *springa* har redaktören i SAOB att under en hatt, i *en* räkka moment osv., beskriva ett ordinnehåll där nutida och äldre betydelsestruktur "krockar" i större eller mindre utsträckning. Här utgör SAOB (enligt min uppfattning) i essentiellare grad än i andra sammanhang ett arteget slag av ordbok.

Det finns skäl att diskutera fördelar och nackdelar med *en* ordbok för en flerhundraårig språkperiod visavi *flera* (om man icke valt en samlad beskrivning av hela den nysvenska perioden, liksom av hela nyhögtytskan osv., vore nog *flera ordböcker* bättre alternativ än *två*). Jag nöjer mig med det rätt självklara påpekandet att en samlad beskrivning ger bättre historiskt sammanhang och en nästan ofrånkomligt sannare historisk beskrivning. Den fördelen må vägas mot den komplicering som en samlad lexikalisk beskrivning av en flerhundraårig språkperiod medför. Svårigheterna att ge olika betydelsestrukturer full rättvisa är också större i den samlade beskrivningen.

### Litteratur

- Allén S., Loman B. & Sigurd B. 1986. *Svenska Akademien och svenska språket*
- Ekbo S. & Loman B. 1965. Vägledning till Svenska Akademiens ordbok. I: *Skrifter utgivna av nämnden för svensk språkvård* 32
- Jonsson H. 1983. Om historisk lexikografi av typ SAOB. I: *Nysvenska studier* 63 (1983) 127-158.
- Svenska Akademiens ordbok* (förkortad SAOB). 1898-

Arnbjørg Hageberg

## Norsk Ordbok - i 1950 eit nybrottsverk, i 1991 ein anakronisme?

Planane for *Norsk Ordbok* går attende til slutten av 20-åra. Fyrste heftet kom i 1950, band I i 1966 og band II i 1978. Ordboka skulle bokføra målvoksteren, og verket syner samspelet mellom målføre og nynorsk skriftspråk, - at dialektord blir "lyfte opp" og brukte i alle litteraturslag. Kombinasjonen dialektordbok/litterær ordbok var noko nytt i Norden. Den lange redigeringstida for 12-bandsverket skaper vanskar: Endringar i ordforråd og språkbruk krev språklege omvurderingar. Redaksjonen må ta utfordringane i elektronalderen og nyttiggjera seg teknologien. Kan det bli eit *både-og*, -både bok og base?

### Eit nynorsk ordboksverk

I 1930 løyvde Stortinget det fyrste tilskotet til eit større ordboksverk som skulle ta opp i seg tilfanget i dei eldre ordbøkene over norsk mål, fylla ut med nyare tilfang og dermed syna den historiske framvoksteren av nynorsken. I Sverige og Danmark hadde utarbeiding og prenting av store nasjonale ordboksverk alt kome langt, i Noreg var likeins arbeidet med *Norsk Riksmålsordbok* i godt gjenge. Tida var komen for innsamling til eit nynorsk ordarkiv og utarbeiding av eit stort, vitenskapleg ordboksverk. I tillegg til ekserpert ordtilfang frå alle slag trykte kjelder skreiv hundretals språkinteresserte, ulønte medarbeidarar rundt i bygd og by opp ord og uttrykk frå eige målføre.

I 1950 kom fyrste heftet av ordboksverket. Det vart heitande *Norsk Ordbok* med undertittelen *Ordbok over det norske folkemålet og det nynorske skriftmålet*. Ordboka var ei samanstypping av munnleg og skriftleg ordtilfang, eit slag kombinasjon av dialektordbok og litterær ordbok. Dette heilt spesielle oppsettet for ei nasjonalordbok var eit nybrottsverk - i det minste i Norden. I planleggingsfase II (1945-1950) var det ein del skepsis til dette uvanlege dobbeltføremålet med ordboksverket, særleg framlagd av den svenske språkforskaren Hjalmar Lindroth<sup>1</sup>. Store kombinerte ordboksverk med tilfang både frå målføra og frå litterære kjelder var elles alt igangsette både i Sveits (*Schweizerisches Idiotikon* 1881-) og i Skottland (*Scottish National Dictionary* 1931-76), men begge stader med ein annan språkhistorisk bakgrunn enn den norske, og med eit oppsett temmeleg ulikt *Norsk Ordbok*.

Nynorsken er eit *ungt* skriftmål samanlikna med t.d. svensk og dansk. I førre hundreåret la Ivar Aasen fram sine storverk<sup>2</sup> og skapte ein ny skriftnormal tufta på norske bygdemål. Det nynorske skriftmålet har etter mange rettskrivingsendringar fått ein skriftnormal som samsvarar med talemålet i bygd og by.

### Den skriftlege målvoksteren

Sigmund Skard var ein av hovudpersonane i den fyrste planleggingsfasen av ordboksverket. Han skreiv heftet *Norsk Ordbok. Historie - Plan - Arbeidsskipnad* (1932). Der nemner han særskilt at den skriftlege målvoksteren ikkje er bokført, og at den planlagde ordboka må syna korleis skriftmålet har tilmåta seg for å løysa alle slag ulike oppgåver. Ordboka må syna "dei finaste skifte og avskyggjingar" (Skard 1932:26), og kvart sitat i ordboka skal gje éi avskygging av ordtydinga. Saman vil sitata då gje eit bilete av korleis eit einskilt ord har vorte teke opp i skriftmålet, og har fått fine avbrigde etter innhald, stemning, stil og litteraturslag. Den ferdige ordartikkelen skal syna att heile denne voksteren i samandrag. Den nye ordboka skal gje eit samla utsyn over "*norsk kulturmaal i heilo*" (same stad).

Arbeidet med ordboka kom i gang, men som ventande var, syntet det seg snart at den opphavlege planen med 2-3 band ikkje heldt. Det nyinnsamla tilfanget var altfor omfattande til at det berre kunne føyast til ei beingrind av eldre ordbøker og ordsamlingar. - Etter krigen vart det så tilsett ein heilt ny redaksjon, som starta redigeringa på nytt frå grunnen av. Alf Hellevik var hovudmannen bak dei nye redaksjonsreglane for verket, som no vart planlagt til 4-5 band (Hellevik 1958:6).

### Målføre og skriftspråk. Dokumentasjon

For eit så ungt skriftmål som nynorsken trongst det eit stort tilfang til rein dokumentasjon for å få fram rikdommen av ord og uttrykk og fylgja utviklinga og voksteren på ulike område. Og det var tale om ein dobbel dokumentasjon: - kvar i landet finst orda, og kvar i litteraturen? Det skulle dertil gjevast eit oversyn over former - både over målføreformer ulike stader i landet og over eldre og nyare skriftformer. Det var viktig å koma i gang med utgjevinga, men krava til dokumentasjon og dertil vissa om at mykje stoff framleis venta på å bli nedskrive, har gjort at vidare innsamling heile tida har gått parallelt med redigeringa.

Sigmund Skard la alt i det før nemnde skriftet *Norsk Ordbok* (Skard 1932:29-32) fram det tilfanget som den gongen fanst på ordet *age*. Det fylte 6 spalter, og det gjev eit godt bilete av kor mykje materiale redaktøren har i hendene når stoffet skal komprimerast og ordboksartikkelen utformast. Ordboksredaktøren i dag har til kvart ord gjennomsnittleg mange fleire setlar å ta stilling til. Redigeringa vert difor stendig meir tidkrevjande, men vonleg vert artiklane tilsvarande betre.

Også Hellevik har fleire gonger peika på at eitt av hovudføremåla med ordboksverket er å syna grunnlaget i talemålet og bruken i skriftmålet, særleg syna korleis dialektord og nylaga ord er tekne i bruk og "odla og dyrka" i skriftspråket, korleis dei "gjennom bibel og salmebok og verk av store diktarar er lyfte ut og opp or sitt opphavlege språklege miljø og har utvikla nye tydingsnyansar og fått ein ny stilklang". (Hellevik 1958:7.)

Måla var sette høgt. Var dei *for høge*? Var tilfanget - særleg dei fyrste redigeringsåra - breitt og representativt nok til å sikra fullgod dokumentasjon? Viser ordartiklane brytinga og samspillet mellom målføra og skriftmålet? Kan vi sjå at nynorske forfattarar hentar ord frå sitt eige heimemål og frimodig brukar dei i bøkene sine, og at ord som ikkje er kjende over heile landet, er stovereine i det nynorske skriftmålet i ulike stilslag på ein heilt annan måte enn tilsvarande ord t.d. i svensk eller dansk?<sup>3</sup> Gjev samanstøypingsmodellen dette ordboksverket dei kvalitetane opphavsmennene drøymde om?



## Omtale og kritikk

Offentleg framsett kritikk skulle seia litt om kor vidt redaksjonen har nådd måla som vart sette. - Utanom obligatoriske meldingar etter fyrste heftet (1950) og fyrste bandet (1966) av *Norsk Ordbok*<sup>4</sup> har det vel berre kome to vektige kritiske innlegg. Det eine er frå Oddvar Nes (Nes 1988), som på fleire punkt stiller seg heller kritisk til normeringa av oppslagsorda i *Norsk Ordbok*. Nes går m.a. inn for ei meir etymologiserande linje når det gjeld normeringa av tidlegare ikkje-ordboksførte målføreord. Artikkelen til Nes og svaret frå underskrivne og ein medredaktør (Bø/Hageberg 1990) skulle klart visa at samanstøypingsmodellen byr på mange problem og få enkle løysingar. Redaksjonen har meint at brukarane lettast finn fram i ordboka når oppslagsorda ligg nær uttalen. Og for dei friviljuge medarbeidarane som har sendt inn ord frå sitt eige målføre, er det noko som heiter attkjenningsverdi. Ved talemålsnær normering skulle det dessutan vera større sjanse for at nynorskbrukarar også skriftleg tek i bruk målføreord som tidlegare har hatt for lita dokumentert utbreiing til at dei er tekne med i rettskrivingsordbøkene. For vi vil at *Norsk Ordbok* også skal vera eit tildriv til at gode målføreord vert brukte både i tale og skrift.

Det er elles særleg ein artikkel av Mattias Tveitane: "Språkbruk, ord og ordbøker" (Tveitane 1964) og svar frå Alf Hellevik (Hellevik 1964) som er av interesse. For Tveitane er det nett dette at *Norsk Ordbok* registrerer den litterære språkbruken - og dermed også mål-dyrkinga - som er det viktigaste. Han etterlyser ei meir utførleg tydingsdifferensiering (t.d. mellom ord som *eldhug*, *brennhug* og *annhug*) i nynorsk litterær ordbok. Dette er ei så noigrann differensiering at Hellevik avviser krava som umoglege - eller utopiske - innanfor dei tidsrammene redaktørane må retta seg etter. Den noverande redaksjonen må seia seg samd med Hellevik. Men samstundes er vi her inne på dei nær sagt daglege frustrasjonane omsynet til omfanget av *Norsk Ordbok* stiller oss overfor. Vi ville så gjerne nytta ut det rike tilfanget og gå endå mykje grundigare inn på ordhistoria. I denne samanhengen vil eg elles visa til redaktør i *Jysk Ordbog*, Viggo Sørensen, som har vore inne på det same i artikkelen "Det tager tid at afdække sammenhænge. Nogle betragtninger over redaktionstempo og kulturhistorisk forpligtelse"<sup>5</sup>.

## Målføre og skriftspråk - samspel og bryting

Nokre døme vil visa korleis redaksjonen så langt og innanfor dei fastlagde rammene har løyst oppgåva å syna samspelet mellom målføra og skriftspråket:

Verbet *folna* er eit godt målføreord som er vidt kjent. *Norsk Riksmålsordbok* har éi linje om ordet, og det er merkt "dial.". *Norsk Ordbok* har godt og vel ei spalte om *folna* og deler opp i fleire tydingar: 'tapa seg, visna, missa fargen, døy bort, kverva'. Dessutan: 'bli maktstolen av redsle, gysa', - og endeleg: 'uvita, dåna'. Ordartikkelen syner med fyldig dokumentasjon at ordet har vid utbreiing i målføra, og det er flittig brukt av diktarar og forfattarar. Ordartikkelen har med skjønnlitterære sitat av 16 ulike forfattarar. Eg kan nemna Henrik Krohn, Vetle Vislie og Kristofer Janson av dei "gamle" landsmålsforfattarane, og av dei vi må kalla "yngre" har vi Olav Duun, Hartvig Kiran og Sigmund Skard. Dessutan finn vi *folna* i sitat frå *Bibelen* og *Nynorsk Salmebok*, og ordet er nytta i omsetjingar, såleis i Arne Garborgs gjendikting av *Odysevskvædet*. Her er nokre av sitata: *no folnar skogen kring vårt land* (*Nynorsk salmebok* 1929 nr. 634,1) / *folna i tidleg haust, dådlaus og rådlaus var Noregs konge vorten* (Vetle Vislie: *Inga frå Varteig* 1924, s. 126) / *Mina stod midt på golve og folna*

*bort; det sat berre litt glo att i kinna* (Olav Duun: *I ungdomen* 1922, s. 117) / *mange* (i tysk fangenskap) *vart bleike eller gråleitte, og det var som dei var nær ved å folna burt* (Olav Hoprekstad: *Frå lærarstriden* 1946, s. 122) / *kor du hev follna .. or! Sjukdom gjer ingen mann fager* (Rasmus Steinsvik: *Tolstoiboki* (omsetjing) 2. utg 1898, s. 102) / *eg skvatt so eg folna all igjenom og låg som ein klumsa* (Olav Aukrust: *Dikt i samling* 1954, s. 84) / *Enok kjende harmhugen sin folna* (Arne Garborg: *Skriftir i Samling* 1921-1922 bd. IV, s. 133) / *slik folnar det når ein vaknar, det ein drømte om* (Olav Duun: *Blind-Anders* 1924, s. 17) / *mørket folnar inn. Og soli stig ny* (Olav Sletto: *Dagning* 1920, s. 80).

Eit anna brukande døme kan vera verbet *fjasla* med likestilt hovudform *fjatla*. Verbet tyder 'pusla, fumla, vimsa, vikla', og artikkelen syner områda for former med *-tl* og *-sl*. Eit sitat av Halvor Sandsdalen, som er frå Seljord i Telemark, lyder: *ho for og fjasla og vaska i utrengsmål* (*Silketørklede* 1955, s. 16), medan Jens Tvedt frå Kvinnherad i Hordaland skriv: (ho) *fjatladd med spøtet* (d e strikkesøyet) (*Hamna-tjonet* 1889, s. 80). Frå Hardanger (også i Hordaland) finn vi sitatet *fjanka og fjatla og ikkje (få) .. arbeidet frå hendene* (Halldor O. Opedal: *Makter og Menneske* 1934, s. 44). Under *fjaslast* står eit Jens Tvedt-sitat til: *ho var so fegi at fingrarne fjatlast og ho kunde mest ikkje finna hekti* (*Haavard Britannius* 1908, s. 301). Trønderforfattaren Olav Berkaak skriv: *berre ut mot slutten fjaslas det* (d e Fadervåret) *for henne ein grann* (i boka - *og du skal lenge leve i landet* 1968, s. 236). Nedover i heile artikkelen ser vi korleis tydingane og målføreformene held seg til sine område.

Som ordboksredaktør må eg vedgå at det gjev meg ei glede kvar gong eg såleis registrerer at eit typisk målføreord er brukt av ein forfattar frå det området vi har ordet nedskrivt frå, ofte til og med i særmerkt form. Materialet vårt *held*, i alle fall eit godt stykke på veg.

## Bibelmål og målføreord

Bibelmålet vårt har vore eit takksamt studieobjekt av mange grunnar. Ikkje minst alle dei ulike omsetjingane har gjeve materiale til språkgranskarar. Såleis er det skrivne ei doktoravhandling om nynorsk bibelmål (Bondevik 1986). I utgreiinga her vil det berre verta høve til litt småplukk for å illustrera kva spaltene i *Norsk Ordbok* kan fortelja oss om bruken av målføreord i *Bibelen*.

Ordet *dust* (merkt "dial." i NRO) finn vi i *Norsk Ordbok* i hefte 1, band II, trykt i 1974. Under IV *dust* tyding 1 'dumbe, støv' er det heile 10 sitat frå *Bibelen*, *Nynorsk Salmebok* og annan religiøs litteratur. Ordet ser ut til å vera fullt ut godteke i religiøst mål: *den som steinen fell på, skal han smuldra til dust* (Mattæus 21,44) / *kast deg i dusti for den nådige Gud* (P.M. Madland omsett O. Hallesby: *Dagleg fornying* 1933, s. 41) / *med song so spring me deg i mot og kysser dusti for din fot* (*Norsk songbok* 1927, 113,2) / *eg vil lata ætti di verta som dusti på jordi* (1. Mosebok 13,16). Bibelsitata er frå 1921-omsetjinga, og *Norsk Ordbok*-redaktøren har ikkje teke med at ordet *dust* vart *skifta ut* i den nye omsetjinga av *Det Nye Testamentet* som kom i 1961. Også *Det Nye Testamentet* 1975 og heile *Bibelen* 1978 har *støv* i staden for *dust*, - eller det er nytta omskrivingar.

Ordet *fysne* med tydingane 'lyst, drift, lengt, kjenslerus, oppgløding' finst ikkje i *Norsk Riksmålsordbok*, men det har vid utbreiing både i målføra og nynorsk litteratur. Vi finn det i ordbøkene til Aasen og Ross. Det er eit brukeleg ord og eit forvitneleg ord med ein heil "ordfamilie" kring seg: *fus* (adj), *fysa* (v), *fysemat*, *fysna* (v), *fysneleg*, *fysnestund* for å nemna nokre. *Norsk Ordbok* viser til samansetjingar som *hemnfysne*, *jordfysne*, *krigsfysne*, *veidefysne*

og ærefysne og har sitat der ordet er brukt - i *Bibelen*, i høgprosa og i poesi, i diktverk av Ivar Aasen, Per Sivle, Olav Aukrust og Tarjei Vesaas. I Paulus' brev til galatarane (5,24) les vi (*Bibelen* frå 1921): *dei som høyrer Kristus Jesus til, dei hev krossfest kjøtet med fysnorne og lysterne*. Ordartikkelen *fysne* syner eit nært samspel mellom målføre og skriftspråk på alle nivå. Men i dei nye bibelomsetjingane er dette "odla" ordet skifta ut. I *Det nye Testamentet* frå 1975 står det *lidenskap* i staden for *fysne*.

Vel det best kjende av alle kapitla i *Bibelen* er juleevangeliet. Norsk/dansk *nogen hyrder .. som var ute på marken og holdt nattevakt over sin hjord* (Lukas 2,8) har i eldre nynorske bibelomsetjingar tilsvarande *nokre hyrdingar .. som laag ute og vakta buskapen*. Men i dei nyaste omsetjingane ser vi at både bokmål og nynorsk har *sauene*, til avløysing for tidlegare *hjord* og *buskap*. Både dette skiftet og det at *dust* er blitt til *støv* og *fysne* til *lidenskap*, må vera gjort av pedagogiske omsyn. Moderne bibelomsetjarar har ikkje vilja ha med ord som ikkje høyrer til daglegtalet og av den grunn kan verka noko gammalmodige. Ordhistoria er dermed ikkje lenger berre filologisk/stilistisk. Ulike omsyn har kome inn, og biletet har vorte meir komplisert. Utviklinga har gått andre vegar enn den planleggjarane av *Norsk Ordbok* tenkte seg.

Tilvisingssystemet i *Norsk Ordbok* opnar for å visa til former i tidlegare og seinare bibel-omsetjingar. Dette tilvisingssystemet har vel redaktørane til no mest brukt for å vera trygg på at tolkinga av bibelsitata er rett, slik vi t.d. ser det under *fullnad: til dess at fullnaden av heidningarne er komen inn* (Romarbrevet 11,25 frå 1921-utgåva) har tilvising til 1961-utgåva: *til dess det er innkome så mange heidningar at talet er fullt*. Etter mitt syn bør *Norsk Ordbok* - nettopp for å syna målbruken og ordhistoria i si fulle breidd - visa målføreord som er "lyfte opp" og brukte i alle stilslag, men òg ta med at somme av desse orda no er tekne ut att or *Bibelen*. Å dokumentera også denne utviklinga vil vera å oppfylle påbodet frå Sigmund Skard om å gje eit samla utsyn "over norsk kulturmaal i heilo" (Skard 1932:26).

### Møte med nytida og elektronalderen

Når eg har nytta ordet *anakronisme* i samband med *Norsk Ordbok* er det i høve til edb-teknologien og ropet på dataføring/databaseredigering og ordbøker i maskinell form. Her deler *Norsk Ordbok* lagnad med andre bandsterke, tradisjonelle, nasjonale ordboksverk.

Etter dei fyrste planane og prognosane skulle *Norsk Ordbok* vore ferdig for lengst. Men med berre to band og to hefte ferdigtrykte no i 1991 er vi svært langt frå målet (12-14 band), og vi kan ikkje lata vera å møta og ta stode til dei nye utfordringane.

Eigenleg burde eg ikkje teke opp dette emnet, eg er ingen ekspert på det eg skal greia ut om. For mange vil det som no fylgjer vera svært elementært. Eg får trøysta meg med at eg ikkje er åleine i flokken av ordboksfolk som framleis ser og høyrer med skrekkblanda fryd på det som vert lagt fram av resultat og vyar av kollegaer som har trengt lenger inn i edb-materien.

Med omsyn til *Norsk Ordbok* og det å gjera seg nytte av edb, så er det opplagt eit stort problem at trykking og utgjeving tok til for over 40 år sidan, og at vi enno har så langt fram. I dag ville vel ikkje ei stor historisk ordbok fått slike rammer og eit slikt redaksjonelt oppsett som det vi har i *Norsk Ordbok*. Det skulle vera nok å visa til Island, der dei i 80-åra valde heilt andre løysingar for sine ordboksverk, skreddarsydde for redigering i database. Føresetnadene for slik redigering er at ordtilfanget blir gjort tilgjengeleg på data. I våre ordsamlingar er berre ein forsvinnande liten del av ordsamlingane dataførte - 12.000 ekserpt

mot 3,2 millionar setlar i arkivskuffene. I 1990 kom planane om dataføring av arkiva ved samlingsinstitutta ved Det historisk-filosofiske fakultetet her i Oslo - det såkalla *dokumentasjonsprosjektet*. Eit reknestykke synte at det ville gå med 107 årsverk berre til samlingane i nynorskavdelinga, og då var ikkje korrektur og anna kontrollesing medrekna. No våren 1991 kjem det melding om at prosjektet for nynorskavdelinga sin del i fyrste omgang vil bli å prøva å få innskrive det grunnmanuskriptet som vart utarbeidd i 30-åra i den fyrste planleggingsfasen av *Norsk Ordbok*<sup>6</sup>. Det er nyttig å få det inn på data, men det vil ikkje gje noko revolusjonerande nytt grunnlag for redigeringa av *Norsk Ordbok*.

Det er elles ikkje slik at redaksjonen i *Norsk Ordbok* til no har oversett alt som heiter edb. Den fyrste dataførte ordboka i Noreg (*Norsk landbruksordbok* 1979) vart utarbeidd i nær tilknytning til vår avdeling, og *Bokmålsordboka* og *Nynorskordboka* (begge 1986) vart utarbeidd og manusa dataførte ved instituttet. Revisjon og ajourføring av desse bøkene skjer sjølvstøtt direkte i databasen. *Norsk Ordbok*-manus blir levert heilt trykkeferdig til forlaget - det siste heftet sett med fotosetjar.

Etter kvart har alle redaktørane fått sin PC. Men det har vist seg å vera langt lettare å få pengar til maskinar enn til menneskehjelp. Vi har ikkje brote skikkeleg gjennom muren og *forlangt* eksperthjelp til å velja programvare og utvikla det programmet vi i vår spesielle situasjon treng for å kunna nyttiggjera oss teknologien optimalt. Eg har inntrykk av at ordboksredaksjonane i dei andre nordiske landa ligg noko lenger framme, særleg då Island og Danmark. (Eit stikkord her er *GestorLEX*.) No vonar vi på eit nærare samarbeid, og at vi på vår side kan få dra nytte av pionerarbeidet som er gjort i grannelanda.

I mellomtida arbeider vi i det daglege ut frå det stadiet kvar einskild redaktør har nådd fram til. Redaktørane har dei dataførte delane av manuset inne på harddisken og søker etter mønster i liknande ord som er redigert. Ved søking kan vi òg få hjelp til effektiv konsekvenskontroll. Dermed sparer vi tid i redigeringa, det blir mindre av inkonsekvensar, og det er lettare å kontrollera at alle redigerer så likt råd er, når det gjeld det ordbokstekniske oppsettet. I fylgje Kolbjørn Heggstad - som er vår veteran i datamaskinell språkbehandling - er ein PC med eit godt tekstbehandlingsprogram ingen dårleg hjelpereiskap for ein ordboksredaktør (Heggstad 1991:8). Men føresetnaden er at vi utnyttar systemet *fullt ut* og ikkje slår oss til ro på eit visst nivå og dermed berre brukar ein flik av det hjelpeapparatet vi skulle ha tilgjengeleg i maskin- og programvare. Ei grundig innføring av ein språkinteressert dataekspert som forstod kva akkurat vår gruppe trong, ville sikkert resultert i ein del mindre underverk i det daglege meir tekniske ordboksarbeidet. - I samband med det førnemnde dokumentasjonsprosjektet får vi tilført ekspertise, som tvillaust etter kvart vil koma heile ordboksverket til gode.

### **Brukarane og edb**

Noko vi kan gjera i ein overgangsfase, er å prøva alt no å ta vare på meir av det tankearbeidet som går føre seg i sorterings- og redigeringsfasen. Til no har mykje av dette berre forsvunne i det stille. Somme av oss redigerer noko romsleg og tek med mange sitat i fyrste manuskriptet, for så å velja ut dei aller beste sitata seinare. Manuset blir nedskore før prenting, og det som blir stroke for ikkje å sprengja alle rammer, blir då borte for alltid. I datatida er det lett å ta vare på ein fylldigare versjon, slik at meir av sorterings- og tankearbeidet på komprimeringsstadiet kan koma til nytte for brukarar som måtte ynskja å gå djupare inn i materialet.

Dei dataførte delane av ordboksmanuset er sjølvstykke koda. Men det kan truleg setjast inn fleire kodar for å leggja til rette for meir allsidig søking, t.d. egne kodar for tradisjonsstoff, ordtak og fyndord m.m. Det trengst fleire inngangar til det verdfulle materialet som ofte har blitt liggjande låst og vanskeleg tilgjengeleg. Vi kan t.d. ta ordartiklane om preposisjonane og dei vanlege verba som *gje*, *gjera*, *gå*, *ha* og *koma*. I store ordbøker er kvar av desse artiklane ei avhandling for seg. Det er desse artiklane det tek det lengst tid å utarbeida, og dei inneheld mest verkeleg gjennomarbeidd stoff. Men det blir sjeldan slått opp i dei. No skulle desse avhandlingane ved hjelp av edb-teknologien kunna gjerast tilgjengelege på heilt nye måtar. Ein oppglødd kollega tok eingong sterkt i og sa at det som til no har vore eit stengt gravkammer, skal bli eit skattkammer, og ordsfattane skal hentast fram i lyset.

### Bok eller skjerm

Eg har brukt ordet *anakronisme*, - eit heller negativt ord. Tidlegare hovudredaktør i *Norsk Ordbok*, Alf Hellevik, var då heller ikkje heilt nøgd med ordvalet mitt. Men spørsmålsteiknet etterpå gjev vel ein mistanke om at det kjem eit nektande svar på spørsmålet som er stilt.

Eg er ikkje åleine når eg trur at boka i god gammaldags forstand har ei framtid. Ikkje ein gong dei aller siste utgåvene av elektroniske lommebøker - som er lettare å bera med seg enn vanlege papireksempplar - ser eg på som verkelege konkurrentar. Det vil alltid finnast mange som ikkje støtt vil sitja og glo på ein flimrande skjerm, men heller vil la augo kvila på sider med rolege teikn i ei fint oppsett bok. Og vi vil vel helst ikkje ha det så travelt at vi støtt må velja det alternativet som gjev maksimal kvantitativ uttelling. Eg vonar dei maskinleselege og dei tradisjonelle informasjonskjeldene i trykte utgåver kan eksistera side om side. Kvar har sine føremoner. Dei maskinleselege vil kunna gje fleire variantar og fyldigare, oppdatert dokumentasjon til dei som treng særleg omfattande opplysningar, som vil ha informasjonen raskt og må ha med det nyaste nye. Teknikken vil når det trengst kunna gje hjelp til å dra ut og sortera på ein blunk materiale som vi etter tradisjonelle metodar ville bruka år til å få oversyn over. Men til tankedelen av redigeringarbeidet i tradisjonell leksikografi vil det alltid vera bruk for menneskehjernen. Heldigvis for alle oss som trivs med ordboksredigering.

Det er langt for oss i redaksjonen i *Norsk Ordbok* å sjå til endes i alfabetet, men vi lit på at det med hjelp av edb-teknologien i åra framover vil gå kortare tid mellom kvar gong vi tek eit nytt vakkert band av *Norsk Ordbok* i handa, og vonleg har vi mange med oss når vi gler oss over alt tilfanget som der er samla og utgreidd.

### Notar

1. Lindroth meinte at det ideelle ville vera tre ordbøker, éi for bokmål, éi for det nynorske litteraturmålet og éi rein dialektordbok (Hellevik 1958:6).
2. Ivar Aasen gav ut *Det norske Folkesprogs Grammatikk* 1848 (*Norsk Grammatikk* 1864), *Ordbog over det norske Folkesprog* 1850 (*Norsk Ordbog* 1873), *Prøver af Landsmaalet i Norge* 1853.
3. Karl-Hampus Dahlstedt drøfter i fleire artiklar folkemålsord i høve til svensk riksspråk (Dahlstedt 1959 og Dahlstedt 1986) og siterer Sara Lidman: "Det är .. inte bara för att vara trogen mot miljön, som jag använder dialekten, utan jag strävar också medvetet att genom den vidga och berika det språk som jag skriver. I bygdemålet finner jag ofta uttryck för betydelser och betydelsenyanser, som jag saknar i riksspråket." (Dahlstedt

1959:107.) Men det store nasjonalordboksverket *Svenska Akademiens ordbok* (SAOB) og likeins *Ordbog over det danske sprog* (ODS) har restriktive, om enn *noko* flytande grenser med omsyn til folkemålsord. For at dialektord skal takast med, må dei vera brukte i kjende riksspråksskrifter i forteljande stil. Folkemålsord i dialektfarga dialog blir til vanleg ikkje tekne med. (Ekbo/Loman 1971:8; ODS 1918:XXIV.)

4. Dag Gundersen melde *Norsk Ordbok*, band I i *Maal og Minne* (Gundersen 1967:148,149). Han peikar m.a. på at ei lang redigeringstid vil koma til å skapa store problem for verket. Vurderinga av einskilde ord og ordtypar endrar seg raskt, særleg i eit ungt litteratur- og kultur mål som nynorsken. Gundersen viser med fleire døme kor snøgt foraldningsprosessen i språkbruken set inn, og at redaksjonen difor stendig vil bli stilt overfor krav om språklege omvurderingar.

5. I *Ord & Sag* (1983 s. 38-47) utgjeve av *Institut for jysk sprog- og kulturforskning* skriv ordboksredaktør Viggo Sørensen at kulturelle samanhengar som går fram av ordboksmaterialet, ofte blir splitta opp av oppsettet i den tradisjonelle ordboka med alfabetiske oppslagsord. Opplysningane finst spreidde på ulike oppslagsord. Krysstilvisingar kan retta opp og syna samanhengar eit stykke på veg. Men tidsnauda - og ordboksoppsettet - hindrar redaktøren i å nytta materialet fullt ut og syna språk- og kulturhistoriske samanhengar slik han eller ho gjerne vil. "At redigere en ordbog er at sætte sig ikke bare mellem to, men midt imellem mange stole." (Sørensen 1983:39.)

6. "Grunnmanuskriptet" utarbeidd i 1930-åra er eit råmanuskript på om lag 13 500 maskinskrivne sider (brei marg og dobbel linjeavstand). Det er ei samanstøyping av ordtilfanget i eldre ordsamlingar og ordbøkene til Aasen og Ross, *Nynorsk etymologisk ordbok* av Alf Torp og også *noko* frå andre ordbøker (m.a. Steinar Schjøtt: *Norsk Ordbok* 1914). Alle definisjonar er overførte til nynorsk (1917-rettskrivinga).

## Litteratur

*Bokmålsordboka* 1986. Oslo.

Bondevik, Jarle: *Studiar i det nynorske bibelmålet*. Doktoravhandling. Universitetet i Bergen 1986.

Bø, Reidar. 1989. Arbeidet med *Norsk Ordbok*. I: *Ord og mål. Festskrift til Magne Rommetveit*. 80-90. Oslo.

Bø, Reidar/Hageberg, Arnbjørg. 1990. Om normering i *Norsk Ordbok*. I: *Maal og Minne*: 79-84. Oslo.

Dahlstedt, Karl Hampus. 1959. Folkmål i rikssvensk prosadiktning. I: *Nysvenska Studier*: 106-168. Lund.

Dahlstedt, Karl Hampus. 1986. Dahlstedts hundra ord. Om folkemåls ord i svenskt riksspråk. I: *Språkvård* nr. 4, 5-16. Stockholm.

Ekbo, Sven/Loman, Bengt. 1971. *Vägledning till Svenska Akademiens ordbok*. Andra upplagan. Stockholm.

Gundersen, Dag. 1967. *Norsk Ordbok* Band I: A - doktrinær. I: *Maal og Minne*: 147-152. Oslo.

Heggstad, Kolbjørn. 1991. Datamaskinen, programvarer og leksikologi. I: *Lexikonord. Leksikografi i Norden*: 7-10. Oslo.

Hellevik, Alf. 1958. *Norsk Ordbok*, eit nasjonalt storverk. I: *Forskningsnytt*: 3. Oslo.

Hellevik, Alf. 1964. Meir om språkbruk og ordbøker. I: *Syn og Segn*: 169-174. Oslo.

Nes, Oddvar. 1988. Om normering i *Norsk Ordbok*. I: *Maal og Minne*: 129-143. Oslo.

*Norsk landbruksordbok*. 1979. Oslo.

*Norsk Ordbok* (NO). 1966-. Oslo.

- Norsk Riksmålsordbok* (NRO). 1937-1956. Oslo.
- Nynorskordboka*. 1986. Oslo.
- Ordbog over det danske Sprog* (ODS). 1919-1954. København.
- Ross, Hans. 1895. *Norsk Ordbog* 1895-1913. Christiania.
- Schjøtt, Steinar. 1914. *Norsk Ordbok*. Oslo
- Schweizerisches Idiotikon. Wörterbuch der schweizerdeutschen Sprache*. Bearbeitet von Friedrich Staub et al., 14 Bände. (1991 bis w). Frauenfeld (Verlag Huber) 1881-.
- Scottish National Dictionary* I-X, 1931-1976. Edinburgh.
- Skard, Sigmund. 1932. *Norsk Ordbok. Historie - Plan - Arbeidsskipnad*. Oslo.
- Svenska Akademiens ordbok* (SAOB). 1898-. Lund.
- Sørensen, Viggo. 1983. Det tager tid at afdække sammenhænge. Nogle betragtninger over redaktionstempo og kulturhistorisk forpligtelse. I: *Ord & Sag*: 38-47. Århus.
- Torp, Alf. 1919: *Nynorsk etymologisk ordbok*. Kristiania.
- Tveitane, Matias. 1964. *Språkbruk, ord og ordbøker*. I: *Syn og Segn*: 162 -168. Oslo.
- Aasen, Ivar. 1873. *Norsk Ordbog*. Christiania.

## Almensproglige informationer i fagordbøger

Der er i de sidste år udarbejdet en række artikler om omfanget og kvaliteten af fagsproglige informationer i almensproglige ordbøger. Derimod har den "modsatte" problemstilling ikke fundet særlig interesse. Denne manglende interesse har sandsynligvis to grunde: Dels har metaleksikografien indtil nu først og fremmest beskæftiget sig med almensprogligt relevante temaer, dels er fagsprogsdiskussionen ikke nået frem til en almen anerkendt forståelse af forholdet mellem almensprog og fagsprog. Her plæderes for et tekstlingvistisk udgangspunkt for fagsprogsforståelsen og for den deraf bestemte lemma- og informationsselektion i fagordbøger.

### 1. Fagsproglige informationer i almensproglige ordbøger

Der foreligger en lang række bidrag, der belyser omfanget og kvaliteten af fagsproglige oplysninger i almene ordbøger. Det er et gennemgående træk i flere af disse arbejder, at de peger på en sand invasion af fagord i almene sprogordbøger i de seneste årtier. Tilsvarende har Joachim Mugdan og jeg i en artikel om behandlingen af lingvistiske termer i almene sprogordbøger og lingvistiske fagleksika påvist, at store almene ordbøger anfører flere lingvistiske termer end egentlige fagordbøger, i en række tilfælde bringer de også en mere adækvat definition. Vores undersøgelser viste, at også de store tosprogsordbøger indeholdt op til 85% af en liste med 351 grammatiske termer, sml. den tabellariske oversigt fra Bergenholtz/Mugdan (1985:11):

	antal behandlede termer (ud af 351)	antal termer med en markering "Gramm.", "Ling.", "Phon." etc.
DUDEN-GWB	87 %	77 %
BROCKHAUS-WAHRIG	92 %*	86 %*
KLAPPENBACH/ STEINITZ	69 %	62 %
WAHRIG	80 %	55 %
HARRAPGER-ENGL	84 %*	73 %*
LANGENSCHIEDT DT-ENGL	80 %	68 %
GROSSES WB DT-RUSS	85 %	66 %
MEYER-ENZYKLOPÄDIE	56 %	40 %
BUSSMANN	59 %	[59 %]

Tabel 1

\* Procentberegningen er ikke basert på en mængde af 351, men på antallet av lemmata i de ordbogsbind, som var udkommet (i 1984).



Ikke i alle tilfælde kan man finde en så bred kodificering af fagord; det er også muligt, at sprogleksikografer er fagligt belastede og derfor i særlig grad medtager fagord fra deres eget fag. Hvis man sammenligner med tal taget fra en hyppighedsliste over forkortelser i Gyldendals elektroniske dansk-tyske ordbog er mængden af eksplicite informationer om fagsprog mere beskedent:

17960 m (maskulin)	28 pol (politik)
17701 f (feminin)	23 relig (religion)
13222 n (neutrum)	23 geol (geologi)
421 sø. (søtvæsen)	21 fot (fotografi)
202 jur (jura)	19 prot (protestantisk)
172 tekn (teknik)	18 økon (økonomi)
160 mil. (militær)	18 mat. (matematik)
157 merk (handelsprog)	18 fys (fysik)
106 fodb (fodbold)	16 bio. (biologi)
63 mus. (musik)	13 kem (kemi)
58 med. (medicin)	13 arkit (arkitektur)
49 edb	10 geogr (geografi)
46 gram. (grammatik)	10 arkæol (arkæologi)
45 fly. (flyvning)	9 rel (religion)
43 gym (gymnastik)	9 ast (astronomi, astronautik)
41 elek (elektricitet)	9 anat (anatomi)
40 tele (telekommunikation)	8 psyk (psykologi)
39 jernb (jernbanevæsen)	8 geom (geometri)
37 typ (typografi)	7 fon (fonetik)
37 teat (teater)	4 tømr (tømrersprog)
37 hist. (historie)	2 fysiolog (fysiologi)
34 kat. (katolsk)	1 myt (mytologi)
28 zo (zoologi)	

Tablet 2

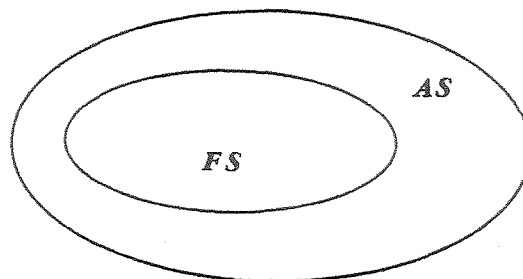
(Mere for en ordens skyld vil jeg gøre opmærksom på, at jeg ikke selv har opfundet den noget usystematiske brug af forkortelsespunktum og uortodokse forkortelser som fx "anat" og "arkit".)

Sådanne faglige markeringer kan uden tvivl være fordelagtige, men personligt er jeg lidt skeptisk over for den trykte ordbog, der - ligesom alle andre ordbøger - ikke ved siden af de diatekniske forkortelser angiver sådanne hyppighedstal eller mere episk i brugervejledningen gør opmærksom på, at mængden af direkte oplysninger om fx matematiske eller fysiologiske termer er meget begrænset. Min skepsis gælder også de store en- og tosprogede ordbøger, der som Duden-GWB anfører en oversigt over ikke mindre end 184 fagsprog, som ordbogen angiveligt behandler. Det skal ikke betvivles, at alle 184 subsprog bliver nævnt et eller andet sted, men at man i de fleste tilfælde ikke får andet end sparsomme oplysninger om subsprogene. For at opnå en mere fuldstændig lemmaselektion forbundet med der dertil hørende informationer foreslår Béjoint (1988:366) at tage udgangspunkt i en fagsystematik. Dette forslag er konsekvent, hvis man mener at kunne opnå en bare rimelig dækning af de mange fag, som almensproglige ordbøger prøver at dække. Det kan man indenfor udvalgte fag, fx sprogvidenskab. Men jeg mener ikke, at denne dækning er opnåelig i fx alle 184 fag, som Duden-GWB lokker sine brugere med. Tværtimod vil jeg her foreslå en tilbageholdende lemmaselektion og informationsselektion, der kun efter meget strenge principper medtager lemmata hhv. oplysninger om fagsproglig brug, som ikke forekommer regelmæssigt i almensproglige tekster.

Med denne indledning er jeg næsten gået for langt ind på det emne, som i modsætning til det af mig valgte allerede er blevet diskuteret i utallige publikationer. Samtidig har jeg allerede brugt nogle af de lingvistiske fagudtryk, som i særlig grad er kontroversielle i fagsprogsforskningen.

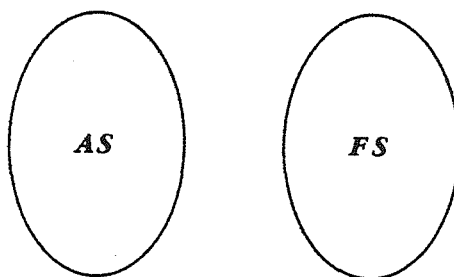
## 2. Almensprog og fagsprog

Diskussionen om en skillelinie mellem fagsprog og almensprog er et tilbagevendende tema i næsten alle almene oversigter over fagsprogsforskningen. Laurén/Nordman, en af de nyere monografier af denne art, vælger at se problemet ud fra forskellige indfaldsvinkler. Den første kan måske kaldes en strukturalistisk systemovervejelse: "Om allmänspråk är det allmäna språket, det som beskrivs i t.ex. en svensk grammatik, är det naturligt att se fackspråket som en del av detta språk" (1987:31). Selvfølgelig er der, siger de, små forskelle i udnyttelsen af det grammatiske systems muligheder i forskellige fagsprog, men stort set kan et fagsprogs grammatik forstås som en ægte delmængde af et almensprogs grammatik. Da Laurén/Nordman bruger grammatik efter et "f.ex.", må dette synspunkt menes at gælde for alle andre eller i det mindste de fleste andre sammenhænge, så der kan tegnes følgende enkle model:



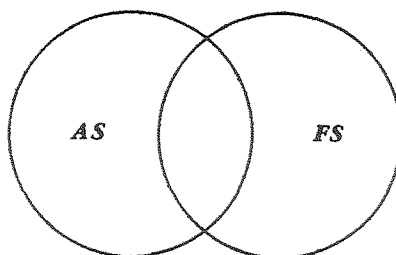
Figur 1

Laurén/Nordman lader det ikke blive stående ved denne sprogsystematiske model, men sætter en anden ved siden af, som de forstår som kommunikationens synsvinkel. Hermed menes, at fagsprog bruges af fagfolk, som forudsætter en vidtgående specialviden om faget. Derimod kan en sådan viden ikke forudsættes i dagligdags samtale, hvorfor der må bruges almensprog for at kunne udveksle tanker. I dette almensprog, siger Laurén/Nordman videre, kan der indgå træk af fagsproglige aspekter, men i en så forenklet form, at helheden ikke mere kan kaldes fagsprog. Fagsprog og almensprog er derfor i denne sammenhæng sidestillede, men også adskilte fænomener (1987:31): "Ur kommunikationens synsvinkel är alltså allmänspråk ock fackspråk sidoställda fenomen." Hermed har vi en ny enkel model:



Figur 2

Der gøres dog opmærksom på, at overgangen mellem fagsprog og almensprog ikke er "entydigt klar". Når fx en forsker giver en populærvidenskabelig fremstilling bruger han ikke mere fagsprog, men måske heller ikke almensprog: "Populariseringen indeholder drag ur fackspråket men liknar allmänspråk mer än fackmannens vanliga texter gör." Ud fra denne argumentation i Laurén/Nordman (1987:32) kan man opstille følgende model:

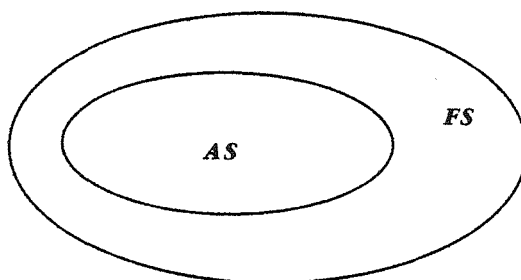


Figur 3

Muligvis ville Laurén/Nordman ikke lægge navn til den sidste model. For selv om det - også efter deres opfattelse - ikke er helt klart, hvor man skal drage grænsen mellem almensprog og fagsprog, bliver antagelsen af en grundlæggende klassifikation med to sprogvarieteter anset

for at være forudsætning for hele fagsprogsdiskussionen, hvor man i hvert tilfælde kan tilordne en tekst til en af varieteterne. I tilfældet med en populariseret fagtekst foreligger der ikke fagsprog, siger Laurén/Nordman (1987:32).

En fjerde mulig grundmodel, hvor almensproget er en delmængde af fagsproget, beskrives af Mikkelsen (1990:2) som den "traditionelle forståelse af forholdet mellem fagsprog og almensprog". Alle de sproglige fænomener, som findes i almensproget, bliver også brugt i fagsprog, som er kendetegnet ved forekomsten af specielle fagord, ord, der ikke bruges i almensproget:



Figur 4

Når man læser fagsproglige bidrag finder man ofte mere komplicerede geometriske modeller, som med en stigende kompleksitet ikke bliver tilsvarende overbevisende, men som de her nævnte fire enkle grundmodeller peger på et grundlæggende problem for fagleksikografien: 1. Kan man sige, at en fagtekst indeholder både fagsproglige og almensproglige elementer? Hvis dette spørgsmål kan besvares med ja, følger heraf næste spørgsmål: 2. Kan og skal en fagordbog give oplysninger om almensproglige elementer i fagtekster? Hvis svaret på første spørgsmål er nej, forbliver det grundlæggende problem: 3. Hvilke af ordene i et bestemt fags tekster bør forklares i en fagordbog?

Svaret på første spørgsmål bliver ofte anset for at være så selvfølgelig, at det ikke bliver stillet. En af de få undtagelser finder man i Ralph 1980 med den retoriske titel: "Hur myckjet fackspråk er fackspråk?". Hans svar er entydigt: meget lidt. I en undersøgelse af arkæologiske tekster med i alt 750.000 tekstord kommer han frem til, at almensproglige leksemer i disse tekster udgør omtrent 97.5 % og fagvokabularet omtrent 2.5 %. Den rent emnespecifikke del af fagvokabularet er endnu mindre, kun 1 % af alle leksemerne i det undersøgte tekstkorpus. Hvis lemmaselektionen til en svensk arkæologisk ordbog, som kun vil indeholde "arkeologiska fackord", skulle foretages ud fra disse overvejelser, ville man kunne nøjes med ca. 500 lemmaer. Jeg er nu ikke helt overbevist om de resultater, som Ralph fremlægger. Dels er optællingen ud fra et korpus på 750.000 tekstord foregået rent manuelt - med den dertil hørende mulighed for unøjagtigheder - dels er den ikke uden videre almenyldig for alle mulige fag. Jeg har fx foretaget en mindre kontroloptælling i det danske genteknologiske korpus og har der foretaget et skøn over antallet af de ordformer, som efter min forståelse af Ralph's opfattelse af fagvokabular kan betegnes som genteknologiske fagord. Af de første 250 leksemer med begyndelsesbogstavet G kan ca. 100 leksemer (40 %) tilordnes det molekylærbiologiske fagvokabular, mens ca. 150 (60 %) ikke har en brug, som er forskellig fra de samme leksemer brug i almensproget. Mængden af komposita i dette

fagsprog er uden tvivl medvirkende til den store fagvokabularandel, sml. følgende uddrag af listen:

fagord	ikke-fagord
gelektroforese	gebærde
gelektroforese-mønster	ged
gelelektroforese	gedde
gelelektroforesbuffer	gedefår
gelfiltrering	gedekrop
gelkoncentration	gedeost
gelkørsel	gehør
gelmasse	geleaktig
geminivirus	gemme
gelretardering	gemyt
gen 1-mærke	general
genaflysning	generalisation
genafskrift	generalisere
genafsnit	generalisering
genaktivitet	generalkonsulat
genalder	generalprøve
genampliktation	generalsekretær

Label 3

Disse tal skal tages med alle mulige forbehold, som spænder videre end de semi-tilhørsforhold, som Ralph regner med. Han har en særlig gruppe, som kaldes "plusord", dvs ord, som uden at være klart bundne til det pågældende fags vokabular er overrepræsenterede i dette fags tekster og som, siger Ralph (1980:167), "indikerar" det pågældende fagområde. Hvis sådanne ord, som fx *datera, fragment, grop, period, rekonstruera, rest, typ, ämne* blev regnet som fagleksemer, ville Ralph komme op på i alt 4 % af alle leksemerne i tekster om arkæologi. Dette tal er stadigvæk langt under de 40 % fagord i genteknologiske tekster, som jeg kom frem til. For mig er det ikke vigtigt, om disse 40 % er helt korrekte eller ej, for tælle-måden vil i alle tilfælde stærkt påvirke resultatet. Det gælder også, når man går ud fra et fælles teoretisk grundlag med en fastlagt metodisk fremgangsmåde. Til et historiografisk værk, som jeg netop har afsluttet i samarbejde med fire tyske kolleger, havde vi et sådant grundlag: "Nur Termini, die explizit genannt und behandelt werden, sind in diesem Register verzeichnet. Pseudoterminologische Häufigkeitsangaben wie *häufig* sind notiert. Allgemeine Begriffe werden nur dann als Stichwort geführt, wenn sie eine Relevanz für die sprachtheoretische Position Heinrich Bauers haben." (Karatas/Meder/Wiechert 1991:119). Hver

medarbejder havde arbejdet med at gennemse et af bindene til Bauers grammatik; i fire tilfælde blev der skrevet mellem 1500 og 2000 terminologikort pr. bind, men i et tilfælde gav gennemlæsningen af et bind kun godt 800 terminologikort. Det er med andre ord næsten en umulighed - selv efter regelmæssige koordinationsmøder - at få forskellige fagfolk til at foretage valg mellem fagord og ikke-fagord med fuldstændig kongruente resultater. Ved lignende tests med udvalgte ord divergerer på tilsvarende vis testpersoners tilordning til et eller flere fagsprog hhv. almensprog i ganske betydelig grad, se Opitz (1979) og Scholfield (1980). Disse iagttagelser svarer i øvrigt til en af konklusionerne i Kalverkämpers ofte citerede grundlæggende bidrag om almensprog og fagsprog (1978:443).

Man skulle måske i stedet prøve at gå den anden vej og prøve at tælle de fælles-sproglige ord. Noget sådant er fx forsøgt af Ruus (1991), som ud fra en undersøgelse af danske korpora med tekster fra romaner, børnebøger, aviser, ugeblade og fagblade prøver at finde frem til de fælles danske ord, de ord, som bliver forstået hhv. brugt af næsten alle danskere. Også i fagbladene forudsætter Ruus (1991:168) at finde fælles danske ord, men er særlig forsigtig: "Hvis et ord er hyppigt i fagteksterne uden at forekomme med sikkerhed i mindst en af de øvrige tekstarter, hører det næppe hjemme blandt de fælles ord." Selvfølgelig er det ikke muligt med sikkerhed at komme frem til helt præcise oplysninger om, hvor mange og hvilke ord danskerne har til fælles. Alligevel er det, siger hun, fristende at give et bud, som hun sætter til at omfatte omkring 1600 "ordbogsord".

Hermed vil jeg ikke sige, at Ruus mener, at man af de måske 20.000 leksemer i det danske genteknologiske tekstkorpus kan trække 1600 fra og anse resten for at være fagord. Men forskellen mellem angivelserne hos Ruus, hos Ralph og sidst fra mit skøn over andelen af ikke-fagord i det genteknologiske korpus er alligevel frapperende. Man fristes til at hælde til den retning indenfor fagsprogsforskningen, som påstår, at alle tekster er fagtekster, at man blot kan tale om en større eller mindre grad af faglighed. Denne opfattelse findes i særlig pointeret grad hos Kalverkämper 1978 og 1980, men er efter min interpretation også grundlæggende for udtrykket LSP, Languages for Special Purposes. Enhver tekst har vel nødvendigvis et specielt formål. Når man med denne indfaldsvinkel betragter alle ord i en fagtekst som tilhørende det pågældende fagsprog, opnår man i alle tilfælde den metodiske fordel, at utilfredsstillende tilordninger ikke er nødvendige: Hvilke af følgende ord tilhører således fagsproget molekylærbiologi?

generation	genforskning
generationsforskel	genforskningspolitik
genetikkomité	genforskningsfelt
genetiklaboratorium	genforsknings sag
genforandring	genforsknings sammenhæng
genforsker	genforskningsudvalg

Tablet 4

Ved min tilordning valgte jeg at regne alle disse ord til ikke-fagvokabularet. Omfanget af denne gruppe er betydeligt; og hvis de alle skulle tilordnes fagvokabularet, ville vi komme frem til en overvægt af leksemer repræsenterede i det genteknologiske korpus, som tilregnes gruppen fagord. En af forklaringerne på den store diskrepans mellem Ralph's resultater og

mine er uden tvivl, at man i høj grad må foretage en graderet skelnen mellem fagsprog med større eller mindre omfang af sproglige udtryk, som kun anvendes faginternt.

I stedet for at begynde på en problematisk grænsedragning mellem fag- og almen-sproglige ord ser den modsatte indfaldsvinkel alle de anførte leksemer som funktionale elementer i en fagsproglig tekst. Om så alle leksemerne med fordel skal lemmatiseres i en ideel fagordbog, er et helt andet spørgsmål. Hvis fx *genflytning* og *genforskning* lemmatiseres, er det næppe påkrævet at forklare *genflytningseksperiment* og *genforskningssammenhæng*. Men det kunne evt. være fordelagtigt at anføre de hyppigst brugte ord i et eksemplarisk udvalg fra typiske fagtekster. Med mindre end 2000 lemmaer vil en fagordbog kunne dække op mod 90 % af alle de ord, der forekommer hhv. skal bruges i en bestemt faggruppes tekster (Hoffmann 1988:4). Muligvis kan en opstilling af grundordforråd i omfattende fagområder som "naturvidenskab" og "samfundsvidenskab", som den er blevet foretaget af Lothar Hoffmann, være til nytte, også som en del af mere specialiserede fagordbøger.

### 3. Principper for lemmaselektion til en fagordbog

Valget af løsning på den diskuterede problematik har umiddelbar indflydelse på lemmaselektionen til en fagordbog. Man må forvente forskellige løsninger til forskellige typer af fagordbøger, men principielt kan der opstilles følgende muligheder:

- 1 kun fagord anføres
- 2 fagord og andre ord, der har en brug, som adskiller sig fra den almensproglige, anføres
- 3 fagord og andre ord, som anses for særligt vigtige i det pågældende fagsprog, anføres
- 4 ord, som anses for at kunne forekomme regelmæssigt i tekster fra et givet fagområde, anføres
- 5 alle ord eller ord med en bestemt mindstefrekvens i en givet mængde af fagtekster anføres

Af disse muligheder undgår kun de to sidste problemet med at skelne mellem fagord og ikke-fagord eller mellem fagord, semifagord og ikke-fagord hhv. almensproglige ord. Til gengæld indebærer de to sidste principper, at en del af problematikken bliver overført til udvælgelsen af tekster til det tekstkorpus, som fremgangsmåden forudsætter.

Men er problemet overhovedet relevant for den fagleksikografiske praksis, i det mindste for den hidtidige praksis? Under en diskussion med ikke-leksikografer nævnte jeg, at fagordbøger ikke kun indeholder fagord, hvilket stødte på stor mistro. Denne mistro er ubegrundet, sml. følgende eksempler fra *Black's Law Dictionary*, som efter sigende gælder for at være et af de amerikanske juridiske standardværker:

**Fall.** One of the four seasons of the year, embracing, in the Northern Hemisphere the three months commencing with the 21st of September and terminating with the 20th of December. Autumn.

**That.** A relative pronoun equivalent to who or which, either singular or plural.

**Window.** An opening made in the wall of a building to admit light and air, and to furnish a view or prospect. The use of this word in law is chiefly in connection with the doctrine of ancient lights and other rights of adjacent owners.

I det mindste ved *window* bliver der givet en henvisning til juridisk sprogbrug. Det gøres dog ikke altid i denne ordbog, sml. fx informationer ved følgende lemmaer: *free, have, make, midsummer-day, the, then, when*. En sådan informationssektion kan også findes i tosprogsfagordbøger, sml. hertil følgende eksempler fra Gullbergs *Svensk-Engelsk Fackordbock. För Näringsliv, Förvaltning, Undervisning och Forskning*:

**kreera** to create

**låta** *hjälppverb* to let, to allow (suffer) ... to, to permit; (*laga att*) to have, to get, to cause; (*formå*) to make; (*låta bli*) to leave (let) alone; *det låt som om* it sounded as if

**yngst** youngest; *-a barnet* [the] youngest child

Nu er det ikke sådan, at man skal lede intensivt for at finde enkelte eksempler med ord, som man uden tvivl også vil kunne finde med en betydning og i en sammenhæng, der ikke væsentligt adskiller sig fra brugen i et af de brede fagområder, Gullberg Sv-E mener at kunne dække. En gennemgang af ialt 18 spalter gav i gennemsnit 7 ord per spalte af typen *yngst*, ord, som man finder i de fleste tekster og som ikke har nogen særlig fagsproglig brug eller betydning. For at få et lidt bredere overblik blev i alt 15 fagordbøger, deriblandt både en-, to og flersprogede fagordbøger, som alle er relevant for den fagsproglige forskning og undervisning i Danmark.

I første omgang kan man læse i ordbøgernes indledning. Af de 15 ordbøger indeholdt seks af dem overvejelser om inddragelsen af ikke-fagsproglige informationer fra det eller de subsprog, der kodificeres i opslagsbogen. Manglende oplysninger om sådanne overvejelser betyder ikke nødvendigvis, at de ikke har fundet sted, hvilket ses af, at én udgave af en ordbog har sådanne oplysninger, en anden ikke (Gubba T-D 1977 hhv. Gubba T-D 1987), eller samme forfatter har med identisk opbyggede ordbøger i det ene tilfælde en sådan information, i det andet ikke (Hauschildt D-F 1988 hhv. Hauschildt D-T). Det er dog påfaldende, at det kun er tosprogsordbøger, der behandler problematikken explicit. Gubba T-D advarer læseren mod manglende informationer om almensproglige betydninger, da det "ofte" kun er den juridiske betydning, der er medtaget; læseren opfordres i stedet til at søge i en almindelig ordbog. Ud fra denne oplysning kan vi indirekte aflede, at der ikke kun er medtaget såkaldte "jurdiske betydninger". Men det er ikke helt klart, om Gubba hermed mener, at man normalt skal bruge to ordbøger ved oversættelse af en tysk juridisk tekst til dansk: fagordbogen + en almensproglig ordbog. Poulsen D-T er i sit forord til første oplag entydig: Denne ordbog tager hensyn til de vanskeligheder, som man måtte få ved oversættelse af danske handelstekster til tysk. Det gælder, "hvad enten det drejer sig om fagsproglige eller almensproglige problemer". Således argumenterer også Ludvigsen E-D 1943, der begrundet denne fremgangsmåde med fordelene ved at nøjes med én ordbog. Unnerup-Madsen D-E (1989), som - uden at angive det - øjensynligt i stort omfang har brugt Poulsen D-T som basis, har ligeledes medtaget det, hun betegner som "almensproglige udtryk og vendinger", men anser det dog for "nødvendigt" at søge forklaringer og oversættelsesmuligheder i "almindelige dansk-tyske ordbøger".

Det er indlysende, at leksikografer, der udarbejder handelsordbøger, i særlig grad ser behovet for at medtage alle ord og betydninger uden at kunne eller ville skelne mellem en brug i handelssprog og en udenfor. I praksis finder man ikke kun en sådan fremgangsmåde i handelsordbøger, men også i praktisk taget alle fagområders opslagsværker, dog snarere i tosprogsordbøger end i ensprogede ordbøger. Der er foretaget en optælling i de anførte



fagordbøger, hvor der er skelnet mellem tre slags ord: Fagord er ord, der ikke kan forstås uden en specialviden om dette fag. Til "tvivlstilfælde" regnes ord, som ganske vist bruges i særlig grad i det pågældende fag, men som alligevel kan forventes forstået af et flertal af sprogets *native speakers*. Til "ikke-fagord" regnes ord, som forekommer i et bredt udsnit af sprogets tekster, som fx romaner, aviser og ugeblade, og som må forventes forstået af næsten alle dette sprogs *native speakers*. Konkret har disse angivelser ført til følgende tilordninger ved gennemgangen af indgangslemmaerne under bogstavet H i den redealfabetisk opbyggede tekniske ordbog Warrern D-E:

**ikke-fagord:** hage, hagl, hale, hamre, harsk, hold-op, horisont, hule, hul, hærværk, hårdknude

**tvivlstilfælde:** hak, halvcirkel, halvkugle, hane, harve, haspe, hovedbestanddel, hugning, hængsel, høvl, håndværksmæssig

**fagord:** halation, heterogen, hexaldehyd, hulendefræser, hyperbolsk stedlinje, hårside

Tilordningen til **ikke-fagord** betyder ikke, at fx *hærværk* ikke kan være et fagord i fx juridisk sprog, men at *hærværk* i teknisk sprog ikke bliver brugt på anden vis end i almensproget. På lignende vis er de øvrige 14 ordbøger blevet gennemset under bogstavet H; som ved Warrern E-D er kun indgangslemmaet blevet talt ved redealfabetisk opbyggede artikler.

Ordbog	Ikke-fagord	Tvivlstilfælde	Fagord
Borum/Popp-Madsen	0 %	14 %	86 %
ECON	42 %	19 %	38 %
von Eyben	0 %	10 %	90 %
Galindo D-Sp	26 %	14 %	61 %
Gubba T-D 1977	13 %	10 %	77 %
Gubba T-D 1987	22 %	7 %	71 %
Hauschildt D-F	61 %	5 %	34 %
Hauschildt D-T	71 %	0 %	29 %
Ludvigsen E-D	71 %	2 %	27 %
Møller-Kristensen	10 %	4 %	86 %
Poulsen D-T	49 %	11 %	40
Unnerup-Madsen D-E	74 %	13 %	5 13 %
Warrern D-E	25 %	14 %	61 %
Woll	2 %	5 %	92 %
Zyphen-Adler	13 %	18 %	64 %

Tabel 5

Et forsøg på en klassificering af ordbøger på grundlag af disse tal og den opdeling i fem klasser, som indledningsvis blev anset for mulige fremgangsmåder i fagordbøger, viser, at Borum/Popp-Madsen, Møller-Kristensen, von Eyben og Woll har brugt princippet fra type 1. Her findes stort set kun lemmata, som entydigt er fagtermer. Om de øvrige ordbøger kan selv om de procentuelle angivelser skal tages med alle mulige forbehold - fastslås, at andelen af ikke-fagsproglige oplysninger er betydelig i en række tosprogede ordbøger. Af disse har Ludvigsen E-D samt Poulsen D-T og måske også Hauschildt D-F, Hauschildt D-T og Unnerup-Madsen D-E gjort brug af princip 4, der anfører alle ord, som antages at forekomme regelmæssigt i tekster fra det pågældende fagområde. Det er ikke umiddelbart muligt at se, efter hvilke kriterier ECON, Galindo D-Sp, Gubba T-D og Warrem D-E har udvalgt informationer af ikke-fagsproglig art. De enkelte eksempler tyder på, at det snarere må være det tredje princip, som er blevet brugt: Fagord og andre ord, som anses for særligt vigtige i det pågældende fagsprog, anføres, sml. eksempler fra den økonomiske ordbog ECON: *Hafen, Hafenarbeiter, halb, (zum) halben Preise, Händler, Häufigkeit, Hotel.*

En endnu større andel af ord, som ikke kan anses for at være fagtermer, indeholder Hoffmann. Denne ordbog lemmatiserer det centrale ordforråd i naturvidenskabelige russiske, engelske og franske tekster. Ved det centrale ordforråd forstår Hoffmann de hyppigste knap 2000 leksemer i et tekstkorpus med naturvidenskabelige tekster. I hver sin ordbogsdel bliver de hhv. russiske, engelske og franske lemmaer forbundet med tyske ækvivalensangivelser. Hvis en af ækvivalensangivelserne er særlig hyppig i en af de tre faggrupper, som Hoffmann tager hensyn til, findes en henvisning til kemi (C), matematik (M) eller fysik (P), sml. følgende artikler i den engelsk-tyske del:

- highly** sehr, stark, höchst  
**hint** hinweisen, andeuten; Hinweis M  
**his** sein *Possessivpronomen*  
**hold (held, held)** halten; enthalten, fassen; tragen; gelten, stimmen; meinen, glauben; halten für  
**hole** Loch; Öffnung; Defektelektron P C

Ordbogen er kendetegnet ved en meget enkel mikrostruktur, der kun indeholder ganske få grammatiske informationer og kollokationsangivelser og ingen eksempler fra de undersøgte tekster. Ud over de fagdifferentierende angivelser som ved **hole** findes desuden - bortset fra brugen af komma hhv. semikolon - ingen hjælp til at adskille de forskellige ækvivalensangivelser fra hinanden. I modsætning hertil indeholder Siliakus' tysk-engelske ordbog over politiske udtryk ved hvert lemma mindst et eksempel. Disse eksempler kan give dels en hjælp til forståelsen af ordets brug, dels nogle implicite informationer om grammatisk brug og om kollokationer:

- Bundestag** <no.m.; -s; -e> (in the FRG:) Parliament (Lower House)  
 In der BRD hat der Bundestag das Recht, Gesetze zu erlassen.  
**Bundestagswahl** <no.f.;--; -en> (federal parliamentary election)  
 Die Wahlen in Niedersachsen werden die bevorstehende Bundestagswahl beeinflussen.

**Bundesversammlung** <no.f.;--;-en> (in der FRG:) Federal Convention; (in Switzerland:) Federal Assembly

Die ursprüngliche Fassung des Art. 60 sah eine Wahl durch die Bundesversammlung (gemeinsame Sitzung von Nationalrat und Bundesrat) vor. (Switzerland)

Mere fordelagtig er eksplicitte oplysninger om kollokationer, som en del ordbøger også bringer. Af de før undersøgte 15 fagordbøger findes sådanne oplysninger i større eller mindre grad i flere af de undersøgte tosproglige fagordbøger. I Poulsen D-T under **konto** bl.a.:

død konto *Sachkonto*  
 fælles konto *Gemeinschaftskonto*  
 føre en konto (*für jmdn*) *ein Konto führen*  
 overtrække en konto *ein Konto überziehen*.

Man kunne måske her - med et vist kendskab til tysk - gætte sig til de verbale forbindelser med *Konto*, men de adjektiviske forbindelser kan man ikke konstruere ud fra et grundkendskab til almensproget. I dette tilfælde drejer det sig om en fagordbog, der i første omgang er udarbejdet for sprogstuderende. Hvis et noget større sprogkendskab hos de potentielle brugere forudsættes, kunne man af pladsbesparende hensyn indskrænke kollokationsangivelser til de tilfælde, hvor internationalismer og ordrette oversættelser svarende til almensprogets sprogbrug ikke kan bruges. Dette er princippet i Kaufmann/Bergenholtz, hvor der under **antistof** ikke findes oplysninger om *monoklonale antistoffer* eller *specifikt antistof*. Derimod kan man i dette fagsprog ikke altid forudsætte, at kendte almensproglige udtryk som *fremstille* (*et antistof*) eller *danne* (*et antistof*) eller *klippe*, *brække*, *skære* (*et DNA*) også virkelig kan oversættes til de direkte engelske udtryk, som også kendes fra almensproget. Netop sådanne kollokationer, som ved første øjekast indeholder almensproglige udtryk, vil efter dette princip i særlig grad være aspiranter til at blive anført under vigtige fagord.

#### 4. Forslag

De gennemførte undersøgelser viser, at der findes en ikke ubetydelig mængde af fagorienterede ordbøger, som ikke kun anfører rene fagtermer. Men det er både teoretisk og praktisk yderst problematisk at betegne disse ikke-fagord som almensproglige ord. En sådan opfattelse måtte medføre, at man skulle kunne opdele en faglig tekst i fagsproglige og almensproglige dele, hvilket ville forudsætte en forældet og ikke hensigtsmæssig forståelse af fagsprog som værende ren terminologi. Det er på den anden side også problematisk at opgive udtrykket almensprog under henvisning til, at alle tekster har et specielt formål eller kun kan beskrives som mere eller mindre fagsproglig. En sådan antagelse er kontraintuitiv, hvis der hermed ikke blot menes, at også ikke-fagtekster er kulturbundne, dvs. at de kun kan forstås med et minimum af kendskab til det pågældende samfund. Antagelsen af et almensprog er tværtimod grundlæggende for hele den gængse fremmedsprogsundervisning og for konceptionen af alle almene grammatiker, hvor der kun lejlighedsvis inddrages fagsproglige aspekter i fremstillingen af den sproglige varietet, som gælder for at være ikke-specialiseret og derfor betegnes som almensprog.

Jeg mener, at begge antagelser kan forenes, hvis man forener de to indfaldsvinkler, der blev beskrevet i Laurén/Nordman (1987:31). Som empirisk basis for fremstillingen i gængse grammatiker og ikke-specialiserede ordbøger foreligger i det optimale tilfælde en tekstmængde, der vil bestå af tekster, som det store flertal regelmæssigt reciperer. Når det drejer sig om skrevne tekster, vil det være aviser, ugeblade og romaner, sml. Bergenholtz (1988). En ordbog eller en grammatik udarbejdet på grundlag af en sådan empirisk basis vil jeg kalde en almensproglig ordbog hhv. grammatik. I praksis bliver en sådan fremgangsmåde i mange tilfælde brugt uden nøjagtige undersøgelser af et tekstkorpus, men den i stedet brugte fremgangsmåde med introspektive angivelser, der forbindes med tilfældige belæg, har det samme mål.

Et fagsprog kan på lignende vis i første omgang bestemmes af en virkelig eller en antaget empirisk basis. Den mængde af tekster, man ekstonialt vil tilordne det pågældende fag, udgør den basis, der danner grundlaget for fagets sproglige beskrivelse. Denne bestemmelse har den fordel, at man kan antage et meget bredt fag eller en mængde af "beslægtede" fag eller en mere indskrænket fagforståelse alt efter, om man ønsker at udarbejde en meget specialiseret fagordbog eller et bredt spektrum af fag, som fx i en teknisk ordbog.

I almensproget findes der uden tvivl ord, som med én begrænsning findes med den samme funktion og betydning i et, flere eller alle fagsprog. Sådanne ord vil jeg kalde **fællesord**. Alt efter, om man sammenligner det almensproglige ordforråd med et, flere eller alle fagsprog, vil mængden af fællesord være større eller mindre. Indskrænkningen er, at der er kollokationelle forskelle, hvilket kan have følger for konceptionen af visse fagordbøger. Tilsvarende kan man tale om **fællesgrammatik**, hvor de grammatiske regler, som er fælles for almensproget og et eller flere fagsprog, kan beskrives. I praksis kan forskellene være mest interessante for fagleksikografien, hvor man som princip for selektion af grammatiske informationer kan vælge at anføre det, som adskiller sig fra almensprogets grammatik eller som ikke er så sikkert forankret deri, at dens beherskelse kan forudsættes (sådan fx Kaufmann/Bergenholtz 1992).

Med disse argumenter er vi nået til det, der også i fagleksikografien er den store ubekendte: brugeren. Man kan opstille følgende grundsituationer, hvor beskæftigelsen med et bestemt fag kan føre til behovet for at bruge en ordbog:

- (a) reception af en fagtekst på modersmålet
- (b) reception af en fremmedsproglig fagtekst
- (c) produktion af en fagtekst på modersmålet
- (d) produktion af en fremmedsproglig fagtekst
- (e) oversættelse fra en fremmedsproglig fagtekst
- (f) oversættelse til en fremmedsproglig fagtekst

Disse grundtyper kan forbindes med de ovenfor opstillede muligheder for inddragelse af informationer om det pågældende fagsprog, som går ud over terminologiske forhold:

- 1 kun fagord anføres
- 2 fagord og andre ord, der har en brug, som adskiller sig fra den almensproglige, anføres
- 3 fagord og andre ord, som anses for særligt vigtige i det pågældende fagsprog, anføres
- 4 ord, som anses for at kunne forekomme regelmæssigt i tekster fra et givet fagområde, anføres
- 5 alle ord eller ord med en bestemt mindstefrekvens i en givet mængde af fagtekster anføres

Til grundtype (a) vil en ensproget fagordbog af type 1, som kun indeholder fagord, være tilstrækkelig. Til (b) må man oftest klare sig med en ensproget fagordbog af type 1. Mere optimal ville en tosproget ordbog af type 2 være, som lemmatiserer fagord og ikke-fagord, som ikke kan betegnes som fællesord. En sådan ordbog skulle dog ud over ækvivalensangivelser også omfatte en faglig forklaring. Til en brugersituation af grundtype (c) foreslås ligeledes en forklarende ensproget fagordbog af type 2 eller 3. Til grundsituation (d) vil en ensproget eller forklarende tosproget ordbog i en kombination af type 2 og 5 være optimal; således vil ordbogsbrugeren finde informationer om alle de hyppigste ord, som han måtte få brug for under tekstproduktionen, men også om alle fagord og ord, som måtte volde særlige problemer. En kombination af type 2 og 4 ville opfylde samme formål, men vil have et mere usikkert empirisk grundlag. Til type (e) vil en forklarende tosproget ordbog af type 2 være ønskelig. Til ordbogssituation (f) vil jeg foreslå en forklarende tosproget ordbog i en kombination af de hyppigste fællesord, alle fagtermini og andre ord, som ikke kan regnes for fællesord på grund af en særlig brug, dvs en kombination af type 2 og 5.

## 5. Slutbemærkning

Titlen på dette bidrag var "almensproglige informationer i fagordbøger". Mere i overensstemmelse med den fremlagte opfattelse ville følgende lidt mindre alment forståelige titel have været: **Informationer om fællesord i fagordbøger**. Den foreslåede selektion af fællesord til fagordbøger under hensyntagen til forskellige grundlæggende brugersituationer er ikke ment som et forslag til at udarbejde et helt urealistisk stort antal ordbøger til alle fag, sprog og sprogpår svarende til hver brugersituation. Tværtimod er det en kendsgerning, at selv om man forestillede sig den ekstreme løsning, at man satte alle verdens lingvister til udelukkende at udarbejde fagordbøger, ville de alligevel næppe kunne udarbejde alle de ordbøger, der hver for sig kan blive brug for i de forskellige brugersituationer. En virkelig løsning på dette dilemma kan efter min mening ikke findes, men en **multifunktional fagordbog**, dvs en forklarende ordbog, der både indeholder fagtermini og de hyppigste fællesord, vil ikke bare kunne bruges ved oversættelse til og fra det fremmede sprog, men også ved reception af modersmålstekster og til friere tekstproduktion på det fremmede sprog.

## Litteratur

- Béjoint, Henri. 1988. Scientific and technical Words in General Dictionaries. I: *International Journal of Lexicography* 1: 354-368.
- Bergenholtz, Henning. 1988. DK87: Et korpus med dansk almensprog. I: *Hermes* 1: 229-237.
- Bergenholtz, Henning/Mugdan, Joachim. 1985. Linguistic Terms in German and English Dictionaries. I: *Lexicographica* 1 : 3-23.
- Bergenholtz, Henning et al. 1991. *Wegweiser durch die Grammatik von Heinrich Bauer. Verzeichnisse und Erläuterungen. Bearbeitet von Henning Bergenholtz, Andreas Dörner, Gregor Meder und Ramona Karatas*. Berlin: de Gruyter.
- Black's Law Dictionary = Henry Cambell Black: *Black's Law Dictionary. Definitions of the Terms and phrases of American and English Jurisprudence, Ancient an Modern*. Fifth Edition by the Publisher's Editorial Staff. Contributing Authors: Joseph R. Nolan and M.J.Connolly. St. Paul Minn: West Publishing Co 1979.

- Borum/Popp-Madsen = O.A.Borum/C.Popp-Madsen: *Juridisk Ordbog*. København: G.E.C. Gads Forlag 1934.
- Duden-GWB = *Duden. Das große Wörterbuch der deutschen Sprache in sechs Bänden*. Hrsg. und bearb. vom Wissenschaftlichen Rat und den Mitarbeitern der Dudenredaktion unter Leitung von Günther Drosdowski. Mannheim/Wien/Zürich: Bibliographisches Institut. Bd. 1: A-Ci 1976, Bd. 2: Ci-F 1976, Bd. 3: G-Kal 1977, Bd. 4: Kam-N 1978, Bd. 5: O-So 1980, Bd. 6: Sp-Z 1981.
- ECON = *ECON-Wirtschaftswörterbuch: Alle wichtigen Begriffe in einem Band*. Düsseldorf/Wien/ New York: ECON Verl. 1989.
- Galindo D-Sp = Orlando Galindo: *Erhvervs- og økonomiordbog Dansk-Spansk*. København: GTO 1989.
- Gubba T-D 1977 = Wilhelm Gubba: *juridisk ordbog tysk-dansk*. Mårslet: W. Gubbas Forlag 1977.
- Gubba T-D 1987 = Wilhelm Gubba: *juridisk ordbog tysk-dansk*. Mårslet: W. Gubbas Forlag 1987.
- Gullberg Sv-E = Ingvar E. Gullberg: *Svensk-Engelsk Fackordbock. För Näringsliv, Förvaltning, Undervisning och Forskning*. Stockholm: P.A.Norstedt och Sönners Forlag 1977.
- Hauschildt D-F = Gurli Hauschildt: *Handelsfaglig ordbog Dansk-Fransk* 2. udg. Horsens: åløkkes handelsfaglige ordbøger 1988.
- Hauschildt D-T = Gurli Hauschildt: *Handelsfaglig ordbog Dansk-Tysk* 2. udg. Horsens: åløkkes handelsfaglige ordbøger 1988.
- Hoffmann = Lothar Hoffmann: *Grundwortschatz Naturwissenschaften. Russisch-Deutsch, Englisch-Deutsch, Französisch-Deutsch*. Leipzig: VEB Verlag Enzyklopädie 1988.
- Kalverkämper, Hartwig. 1978. Die Problematik von Fachsprache und Gemeinsprache; in: *Sprachwissenschaft* 3: 406-444.
- Kalverkämper, Hartwig. 1980. Die Axiomatik der Fachsprachen-Forschung. I: *Fachsprache* 2: 2-20.
- Karatas, Ramona/Meder, Gregor/Wiechert, Ingrid. 1991. Die Anlage der Indizes. I: Bergenholtz et al.: 117-122.
- Kaufmann/Bergenholtz = Uwe Kaufmann/Henning Bergenholtz: *Genteknologisk ordbog. Dansk-engelsk/engelsk-dansk molekylærbiologi og DNA-teknologi*. København: Gad 1992 (under forberedelse).
- Laurén, Christer/Nordman, Marianne. 1987. *Från kunskapens frukt till Babels Torn. En bok om fackspråk*. Stockholm: Liber.
- Ludvigsen E-D = Vilhelm Ludvigsen: *Engelsk Handelsordbog 1. Del Engelsk-Dansk* Revideret Udgave. Steen Hasselbachs Forlag 1943.
- Mikkelsen, Hans Kristian. 1990. Fagsprog og almensprog: særsprog og fællessprog. I: *LSP Nyhedsbrev* 2: 1-6.
- Møller-Kristensen = Svend Møller-Kristensen / Erik Nielsen / Henning Nielsen: *Erhvervsfaglig Ordbog*. Risskov: Forlaget MERCIO 1984.
- Opitz, K. 1979. Technical Dictionaries: Testing the Requirements of the Professional User. In: Reinhard R.K. Hartmann (Ed.): *Dictionaries and Their User: Proceedings of the 1978 BAAL Seminar on Lexicography*. (= *Exeter Linguistic Studies* 4). Exeter: University of Exeter: 89-95.

- Pongrácz, Judit. 1987. Zur Rolle der allgemeinsprachlichen Lexik in der Fachkommunikation; in: *Fachsprache und Allgemeinsprache*, hrsg. von Sabine Dallmann. Budapest: Lektorat für deutsche Sprache und Literatur: 5-21. (= *Unser Thema* 4).
- Poulsen D-T = Sven-Olaf Poulsen: *Dansk-Tysk Handelsordbog*. Harlev: A. M. Kjærs Forlag 1987.
- Ralph, Bo. 1980. Hur mycket fackspråk är fackspråk? I: *Svenskans beskrivning 12. Förhandlingar vid Sammankomst för att dryfta frågor rörande svenskans beskrivning 12*. Utgivna av Sigurd Fries och Claes-Christian Elert. Umeå: Universitetet i Umeå: 163-172.
- Ruus, Hanne. 1991. Fælles danske ord; in: *Auditorium X. Dansk før, nu - og i fremtiden? En antologi om moderne sprogforskning*, red. af Erik Hansen, Inge Lise Petersen, Ib Poulsen. [København]: Amanda : 158-180.
- Scholfield, P.J. 1980. Evaluating Selection Policy and Grammatical and Semantic Information in an EST Dictionary. In: *Fachsprache 2*: 98-108.
- Siliakus T-E = H.J.Siliakus: *500 German Political Terms and their English Translations together with 500 useful Contexts*. Adelaide: University of Adelaide 1986.
- Unnerup-Madsen D-E = Dorthe Unnerup-Madsen: *Dansk-Engelsk Handelsordbog "For many Reasons" med hovedvægt på forholdsordsforbindelser* 3.udgave, 2. oplag. Holstebro: Venus 1989.
- von Eyben = W.E. von Eyben: *Juridisk Ordbog* syvende udgave. København: G.E.C. Gads Forlag 1987.
- Warrern D-E = Allan Warrern; *Dansk-Engelsk Teknisk Ordbog* 6.udgave. København: Clausen 1986.
- Woll = Artur Woll: *Wirtschaftslexikon* dritte, vollständig überarbeitete und erweiterte Auflage. München/Wien: R. Oldenburg Verlag 1988.
- Zyphen-Adeler = H.Zyphen-Adeler: *Forretningsordbogen — Handelsmæssige, børs-, bank-, forsikrings-, transport-, told-, bogholderi- og revisionstekniske samt juridiske Udtryk og Vendinger, Mønt, Maal, Vægt m.m.* København: Gyldendalske Boghandel/Nordisk Forlag 1944.

## Brikker til en typologi over danske fagordbøger

Målet med artiklen er at finde kriterier til brug ved en typologisering af fagordbøger. Det foreslås at skelne mellem tre typologiske dimensioner: en materiel, en tekstuel og en funktionel, svarende til Hallidays tre basale kommunikationsaspekter: det ideationelle, det tekstuelle og det interpersonelle aspekt. Inden for hver af disse dimensioner hævdes tre parametre at være særligt vigtige, og af kapitel 4 fremgår det, at kombinationer/korrelationer af træk fra de i alt ni parametre kan give det ønskede typologiseringsgrundlag.

### 1. Indledning

De overvejelser og resultater, der fremlægges i denne artikel, indgår i et større forskningsprojekt, som jeg er ansat til at udføre. Projektet har som mål at etablere ny viden om tosprogede fagordbøger, en viden som kan være med til at danne grundlaget for fremtidige fagordbøger af interesse for en dansk bruger. Det siger sig selv, at det med denne målsætning ligger lige for at fokusere på ordbøger, der har dansk som udgangssprog eller målsprog.

Mit forskningsprojekt rummer flg. hovedelementer: 1) En status over danske fagordbøger. 2) Forskning i og udvikling af teorien bag tosprogede fagordbøger til oversættelse under hensyntagen til forskningsresultaterne fra den almene (meta)leksikografi og fagsprogsforskningen generelt. 3) Udarbejdelse af koncepter til tosprogede oversættelsesordbøger inden for nogle udvalgte fagområder.

Denne artikel handler altså især om det første punkt. Jeg skal med det samme understrege, at jeg befinder mig i begyndelsen af opgaven, og at de resultater, som her fremlægges, i høj grad er behæftet med usikkerhed. Dog finder jeg, at mine foreløbige undersøgelser kan udstikke retningslinjer for en typologisk oversigt over eksisterende fagordbøger.

### 2. Generelt om danske fagordbøger

Indtil nu har jeg ved gennemsyn af diverse bibliografier samt kartoteker og databaser ved danske forskningsinstitutioner samlet ca. 1400 titler og lagt dem ind i en bibliografisk database. Jeg har samtidig registreret alle relevante oplysninger, som kunne hentes ud fra bøgernes titler og de oplysninger, som bibliografierne, henholdsvis kartotekerne og databaserne, kunne give. Imidlertid kan man ikke uden videre stole på, at en ordbog nu også er en ordbog. Foruden de rent metaleksikografiske overvejelser om, hvad man overhovedet skal forstå ved en "ordbog", kan det volde meget store problemer, hvis man satses på at finde alle fagordbøger frem fra mængden af bøger, som ikke med nogen ret kan kaldes ved dette navn. På den ene side synes selve betegnelsen "ordbog" at have positive konnotationer. Rammerne for hvad der kaldes ved dette navn er derfor vide. På den anden side er der en



mængde andre ord og udtryk, der kan bruges til at dække det begreb, som ordet "ordbog" også dækker. Jeg kan i flæng remse flg. op:

"1090 engelske økonomiudtryk", album, begreber, benævnelser, betegnelser, "bevingede ord", "Ca. 1000 navne", "dansk-engelsk-fransk-tysk", definitioner, dictionnaire, encyklopædi, fagudtryk, faktaordbog, folkenavne, forkortelser, fortegnelse, "fra A til Z", fremmedord, gloser, glossar(ium), "hvem-hvad-hvor", "hvordan udtales...?", håndbog, index, katalog, konkordans, leksikon, liste, minilex, navne, "navne på stater", "nogle geografiske navnes form, udtale og betydning", nomenklatur, "operabogen", opslagsbog, ord, ordbog (og div. sammensætninger såsom fagordbog, handelsordbog, handelsskoleordbog, teknisk ordbog, fremmedordbog, realordbog), ordforklaringer, ordliste, ordnøgle, ordsprog, "oversættelse af udtryk", parlør, register, registrant, standard, standardudtryk, "systematisk skriftstedssamling", termer, terminologi, terminologiliste, thesaurus, udtryk, vejviser, vokabularium.

Ikke alle disse ord står nødvendigvis for indholdet "(fag)ordbog", men muligheden er der, og på det indledende stadium gør man nok bedst i at tage hensyn til i hvert fald de fleste af disse "mistænkelige" ord.

## 2.1 UDK-grupperne

Til inddelingen af ordbøgerne efter fag benytter jeg mig af det universelle decimalklassifikationssystem, UDK, som bruges til klassifikation på mange biblioteker verden over. Vha. tallene 0-9 kan alt på denne jord få sin plads i systemet - og så er tallet 4 endda ikke udnyttet endnu. Der er ganske vist forskelle fra land til land, ja fra bibliotek til bibliotek, når det gælder specificeringen ud i de yderste decimaler. Men hvis vi holder os til de mest generelle klasser, vil der være sammenfald også på tværs af landegrænser. De 9 udnyttede hovedklasser er flg.:

- |   |  |
|---|--|
| 0 | Videnskab og kultur, alment                |
| 1 | Filosofi                                   |
| 2 | Religion                                   |
| 3 | Samfundsvidenskab                          |
| 5 | Matematik og naturvidenskab, alment        |
| 6 | Anvendte videnskaber                       |
| 7 | Fysisk planlægning. Kunst. Arkitektur m.m. |
| 8 | Sprog- og litteraturvidenskab              |
| 9 | Geografi. Biografi. Historie               |

Inden jeg kommer nærmere ind på mit hovedtema, nemlig de typologiske træk ved fagordbøgerne, vil jeg nævne nogle tendenser i fordelingen af fagordbøger på de forskellige faggrupper. Jeg ønsker ikke at sætte eksakte tal på endnu, eftersom den endelige opgørelse først kan foretages, når jeg har arbejdet hele materialet igennem.

## 2.2 Hovedtendenser

Med de nævnte forbehold, vil jeg driste mig til at angive nogle omtrentlige procentsatser for, hvordan fagordbøgerne, foreløbig blot defineret som det, der ikke er almensproglige ordbøger, fordeler sig på de ni brede fagområder.

0:	2,5 %
1:	2,5 %
3:	18,5 %
5:	9 %
6:	51,5 %
7:	8 %
8:	2 %
9:	2 %

Altså tegner de anvendte naturvidenskaber, teknikken i bred forstand, sig alene for halvdelen af fagordbøger, mens samfundsvidenskaberne tager sig af størstedelen af resten.

Fordelingen af ensprogede og ikke-ensprogede fagordbøger inden for de samme grupper viser at de samme to hovedområder adskiller sig påfaldende fra alle de øvrige ved, at de ikke-ensprogede fagordbøger er i klart flertal (område 0 har dog også et lille flertal af ikke-ensprogede fagordbøger). Behovet for at kunne forstå, henholdsvis gøre sig forståelig på et andet sprog end dansk, synes således klart størst inden for samfundsvidenskaberne og de anvendte naturvidenskaber. Sammenholdt med hvad det er for nogle fag og især fagsprog, der undervises i på de danske handelshøjskoleers erhvervsproglige fakulteter, nemlig jura, økonomi og teknik, kan der ikke siges at være noget overraskende i denne konstatering.

Hvis vi videre ser på, hvordan de ikke-ensprogede ordbøger fordeler sig på to- og flersprogede fagordbøger, ser vi, måske overraskende, at både økonomi og jura har færre tosprogede end flersprogede fagordbøger, mens omvendt de tekniske fag har et flertal af tosprogede ordbøger.

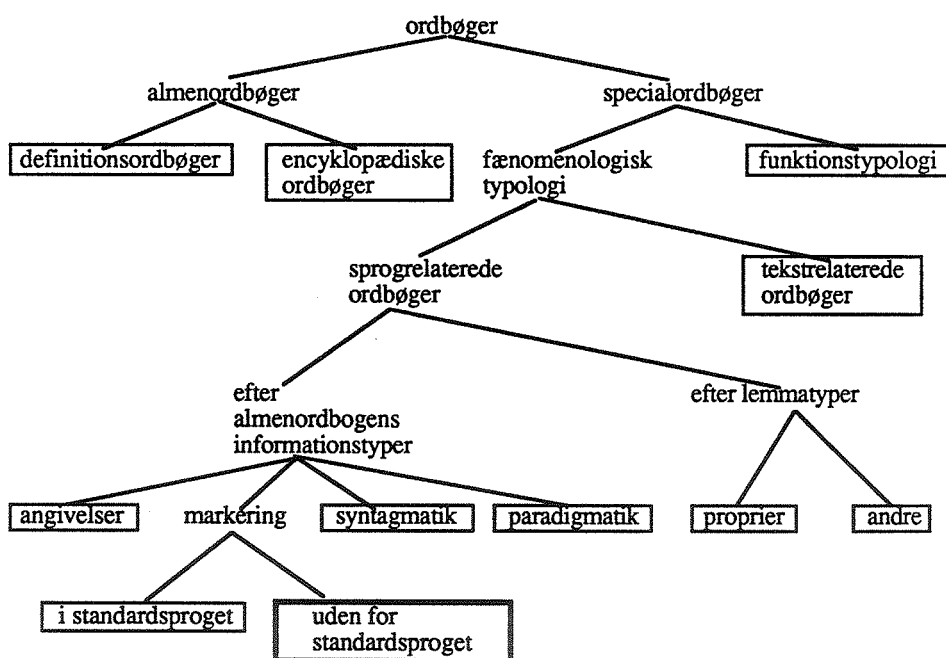
## 3. Typologiske træk

Oven på disse rent kvantitative opgørelser vil jeg nu give mig i kast med nogle mere kvalitative overvejelser vedrørende typologiseringen af fagordbøgerne. Det gælder i endnu højere grad her, at man må tage hver enkelt fagordbog i øjesyn for at kunne beskrive, hvilken type den hører til. Jeg vil i det følgende skitsere et pilotprojekt, hvor jeg har noteret mig nogle karakteristiske sider af et tilfældigt valgt udsnit af de registrerede potentielle fagordbøger. Målet med denne undersøgelse var at nå frem til nogle kategorier, der kan bruges til beskrivelsen af samtlige fagordbøger. Resten af kapitel 3 redegør for de kriterier, jeg er gået ud fra ved gennemgangen af de knapt 70 udvalgte bøger.

Som tidligere nævnt er et af problemerne, hvad man overhovedet skal forstå ved en "ordbog". At der så i dette tilfælde ikke blot er tale om en "ordbog", men en "fagordbog" gør ikke tingene nemmere. Dermed må man nemlig uvægerligt gøre sig forestillinger om, hvad fagsprog overhovedet er. Men først en kort oversigt over, hvad jeg mener er relevant at inddrage ved typologisering.

I den metaleksikografiske litteratur findes der mange forskellige måder at gruppere ordbøger på. Fx har Hausmann (1989) udarbejdet en generel model til typologisering af

ordbøger (se nedenfor). Modellen viser sin styrke ved at kunne rumme alle de forskellige ordbogstyper, som beskrives i Hausmann/Reichmann/Wiegand/Zgusta 1989, kap. VI-XV. Hausmann kalder selv sin typologi for "fænomenologisk", selvom det ret beset kun er en del af typerne, der kan siges at være udledt på dette grundlag. (Med fænomenologisk menes her, at det er arten af genstandsområdet (tekstrelateret eller sproglateret) og artiklens elementer (lemmata, artikelposter), der gøres til udgangspunkt for en klassifikation af ordbøger.) Hausmanns model er beregnet til at kunne indeholde alle tænkelige ordbøger, og vægten på de fænomenologiske træk er i den henseende rigtig. Der indgår dog også andre kriterier i modellen, netop fordi en rent fænomenologisk tilgang ville betyde en ignorering af centrale typer, deriblandt sagordbøger, fagordbøger, læreordbøger og oversættelsesordbøger. Følgelig er modellen suppleret med nogle kriterier af ikke-fænomenologisk karakter for at tage højde for sådanne mangler (se figur 1).



Figur 1

Når det gælder en model for typologisering af fagordbøger, finder jeg det bedre at lægge vægten på andre sider end de fænomenologiske. Her vil jeg koncentrere mig om tre fundamentalt forskellige perspektiver ved typologisk klassifikation. De er dog ikke mere forskellige, end at de også står i en vis indbyrdes sammenhæng. Denne sammenhæng viser sig ved, at de tre kriterier tilsammen dækker alle aspekter af menneskets kommunikation. De tre elementer i kommunikationsprocessen, som jeg har i tankerne, er dels den **virkelighed**, som kommunikationen primært forholder sig til, dels den **tekst**, der repræsenterer virkeligheden og dels endelig den **situation**, som udgør rammerne for kommunikationen. Intet af disse tre grundelementer kan mangle. Min tese er, at de parametre og de parametertræk, jeg fremhæver i 3.1-3.3., er centrale, når det gælder typologisering af fagordbøger. Det

betyder ikke, at man ikke kan finde andre relevante parametre og parametertræk, men at nedenstående er tilstrækkelige til en overskuelig klassifikation.

### 3.1 Materiel typologi - og afgrænsning af "sprog/sag" samt "almensprog/fagsprog"

Mht. den virkelighed, som er repræsenteret i fagordbøgerne, finder jeg, at der er tre vigtige parametre: antal sprog, virkelighedens art og virkelighedens specialiseringsgrad. "Virkelighedskriterierne" spiller kun en mindre rolle i Hausmanns opstilling. Antal sprog er slet ikke med, idet modellen trods sin store rummelighed kun anvendes til klassifikation af monolingvale ordbøger. Virkelighedens art indgår i modellen for at adskille definitionsordbøger fra encyklopædiske ordbøger, mens de rene sagordbøger (encyklopædier) dog indledningsvist er sorteret fra. Virkelighedens specialiseringsgrad indgår i underinddelingen af det fænomenologiske træk "markering" i form af en distinktion mellem det, der indgår i standardsproget, og det, der falder udenfor. Den første parameter lader sig naturligt inddele i flg.

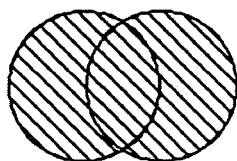
- 1 sprog (monolingvale)
- 2 sprog (bilingvale)
- flere end 2 sprog (polylingvale)
- (= materiel ordbogstypologi, 1. parameter)

De to øvrige materielle parametre kræver lidt mere forklaring pga. deres mindre betydning for den hidtidige metaleksikografiske forskning. Dette gælder især virkelighedens specialiseringsgrad, dvs. forholdet mellem fagsprog og almensprog. Dette spørgsmål er aktuelt inden for fagsprogsforskningen og bliver derfor også et kernepunkt inden for fagleksikografien. Til gengæld har forholdet mellem sag og sprog været vendt og drejet flere gange, bla. ganske udførligt i Hupka 1989. Jeg skal derfor ikke gentage diskussionen. På den anden side er det nødvendigt at tage stilling til begge spørgsmål, ikke mindst når det drejer sig om fagordbøger. Lad mig derfor fremstille mit syn på de to kardinalspørgsmål.

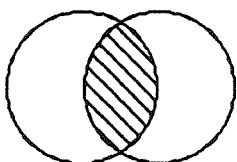
Hvis vi forudsætter, at der findes en virkelighed uden for sproget, siger vi samtidig, at sproget og de sagsforhold, det kan referere til, er, om ikke uafhængige af hinanden, så dog i det mindste ikke identiske. Der er efter min opfattelse tale om **to forskellige dele af en og samme verden eller univers**. Skellet mellem dem er ikke absolut: de overlapper hinanden: "begrebet" udgør deres fællesmængde. Dette er illustreret i figur 2.

Både sag og sprog kan afspejles i ordbogens sproglige virkelighed. Sproget er, som bla. Pragerlingvisterne fremhævede, i stand til at vende sig mod sig selv, at gøre sig selv til genstand for kommunikation (den såkaldt "metasproglige" funktion). Opdelingen i sagordbøger og sprogordbøger er rigtig, hvis man går ud fra den nævnte skelnen mellem sag og sprog. Men også sagordbogen og sprogordbogen overlapper hinanden: definitionen (af begrebet) kan lige så vel optræde som et ledsagende element i sagordbogen som i sprogordbogen - og hvad mere er: det kan være det centrale element i det, som jeg vil kalde "begrebsdefinitionsordbogen".

Hvis vi nu kalder sagens og sprogets differensmængder samt deres fællesmængde for henholdsvis Sp(rog), Be(greb) og Sa(g), får vi flg. syv kombinationsmuligheder: Sp, Be, Sa, SpBe, BeSa, SpSa og SpBeSa. Antallet af kombinationsmuligheder stiger, hvis vi går ud fra, at der i kombinationerne altid er tale om, at den ene af komponenterne har førsteprioritet.



foreningsmængden af sag og sprog:  
ET UNIVERS



fællesmængden af sag og sprog:  
BEGREBER

Figur 2

Altså får vi fig. 12 mulige sammensætninger, hvor den førstnævnte komponent har førsteprioritet, mens andre har andenprioritet:

<i>Sprogordbøger</i>	<i>Begrebsdefinitionsordbøger</i>	<i>Sagordbøger</i>
Sp	Be	Sa
SpBe	BeSp	SaBe
SpSa	BeSa	SaSp
SpBeSa(=SpSaBe)	BeSpSa(=BeSaSp)	SaBeSp(=SaSpBe)

(= materiel ordbogstypologi, 2. parameter)

Netop denne prioritering gør, at billedet ikke bliver helt uoverskueligt. Der er altså altid tale om, at en fagordbog kan placeres som en sprogordbog, en begrebsdefinitionsordbog eller en sagordbog. (Eksempler anføres i 4.)

Skellet mellem sprog-, begrebsdefinitions- og sagordbøger gælder ordbøger generelt. De fleste fagordbøger fordeler sig på de tre nævnte almene hovedtyper, men der er endnu et par stykker repræsenteret i mit udvalg, nemlig:

etymologisk fagordbog  
forkortelsesordbog  
thesaurus  
sætnings- og billedordbog

Den anden distinktion, som er fundamental for fagleksikografien, er den mellem fagsprog og almensprog (= hverdagsprog). I den almensproglige leksikografi spiller denne distinktion en underordnet rolle, eftersom fagordbøger, som allerede vist, placeres langt nede i det fænomenologiske typologiske træ.

Almensproget er, efter min opfattelse, den sproglige del af **den gennemsnitlige sprogudøvers univers**. Den ene halvdel af den gennemsnitlige sprogudøvers univers er den sproglige del. Den anden halvdel af dette univers udgøres som vist ovenfor af de saglige (eller "encyklopædiske") forhold, som den gennemsnitlige sprogudøver har et fortroligt kendskab til. Man kunne evt. kalde denne halvdel for "almensagen".

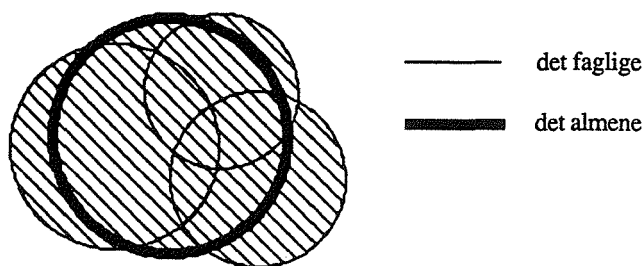
Fagsprog er på tilsvarende vis en halvdel af **et univers, som den gennemsnitlige sprogudøver ikke har fuldt kendskab til**, eller som Henning Spang-Hanssen (1983) ville udtrykke det: fuld sproglig autoritet over. Den anden halvdel af det faglige univers udgøres af de faglige sagsforhold, som også rækker ud over det, som gennemsnitsborgeren er fortrolig med. Der er ikke nødvendigvis noget endeligt antal af faglige universer, og de kan lappe mere eller mindre ind over det almene univers. Det totale univers, og dermed også dets komponenter: de totale sagsforhold og det totale sprog, kan derfor heller ikke afgrænses absolut, men skal forstås som summen af det almene og det faglige univers. Altså:

totalunivers = almenunivers + fagunivers

totalsag = almensag + fagsag

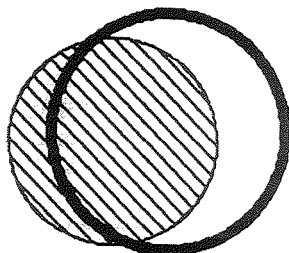
totalsprog = almensprog + fagsprog

Den grafiske fremstilling af forholdet total : almen : faglig viser i figur 3 den hierarkiske relation: det totale univers (skraveret) indeholder det almene og de faglige universer, som indbyrdes er på samme plan.

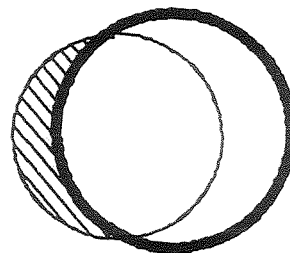


Figur 3

Hvis vi opfatter fagsprog, sådan som jeg fremstiller det i figur 3, kan vi, når vi vil beskæftige os med et fagsprog, vælge at omfatte **hele** fagsproget (skraveret), incl. den del, der er fælles med almensproget (a), eller vi kan vælge kun at behandle den del af fagsproget (skraveret), som falder uden for almensproget (b). Se figur 4.



Figur 4a



Figur 4b

I fagleksikografisk sammenhæng er distinktionen mellem **hele** fagsproget og den specifikke **del** relevant for selektionen af materiale til ordbogen<sup>1</sup>. Således er det karakteristisk for den

del af fagordbøgerne, som man almindeligvis kalder terminologiordbøger, at de kun interesserer sig for den særsproglige del, og inden for denne ofte endda kun for et begrænset udsnit.

Så meget om de to oppositioner, som er helt grundlæggende for leksikografien i almindelighed og for fagleksikografien i særdeleshed.

I min undersøgelse af fagordbøger har jeg fundet det relevant at arbejde med flg. fagsprogsaspekter:

begreber  
 deskriptorer  
 fagsprog  
 navne  
 navne og begreber  
 terminologi  
 (= materiel ordbogstypologi, 3. parameter)

### 3.2 Strukturel typologi

Den strukturelle typologi, som sammen med den omtalte materielle typologi udgør det, som Hausmann kalder den fænomenologiske typologi, tager udgangspunkt i ordbogen som tekst. Man kan skelne tre niveauer: ordbogens struktur (megastruktur), ordlistens struktur (makrostruktur) og artiklens struktur (mikrostruktur).

På det megastrukturelle niveau har jeg valgt at se på, hvilke og tildels hvor mange **komponenter** der optræder. Den eneste megakomponent, man finder i næsten alle ordbøger, er makrostrukturen/ordlisten. Der er dog en undtagelse. Når der blandt kategorierne nedenfor forekommer komponenten "almindelig bog" skyldes det den tidligere omtalte (se kap. 2 ovenfor) uklarhed i afgrænsningen af ordbogsbegrebet. Uklarheden viser sig dels som et praktisk, dels som et teoretisk problem. På den ene side må man regne med, at der som følge af den meget brede titelbaserede afgrænsning af kildematerialet følger en del "støj" med. Noget af materialet vil mao. blive kasseret, fordi det efter nærmere undersøgelse ikke kan regnes med til kategorien "ordbog". En bog, der er struktureret fuldstændig som en almindelig bog, dvs. som en kontinuerlig kohærent tekst, vil ikke kunne opfylde de teoretiske krav til en ordbog. På den anden side forudsætter en sådan udskilning, at der er opstillet nogle kriterier for, hvad der kan betragtes som en ordbog. Jeg har i denne undersøgelse ikke lagt mig fast på en teoretisk definition, og det er bla. grunden til tilstedeværelsen af den her omtalte komponent. Om den teoretiske definition af begrebet ordbog vil jeg kun sige, at det nok vil være hensigtsmæssigt at operere med en prototypisk organisering af det. Der vil således snarere være tale om, at en given bog er en ordbog i større eller mindre grad, end at en bog enten er en ordbog eller ikke er en ordbog. Det er også en af grundene til, at jeg ikke uden videre udelukker, at en bog uden en makrostrukturelt betinget non-kontinuerlig tekststruktur i en eller anden forstand kan være en ordbog. Dog skal det nævnes, at en bog med "almindelig bog"-struktur naturligvis ikke lader sig analysere på makrostrukturplanet, eftersom det er karakteriseret ved ikke at have nogen makrostruktur. En sådan bog kan heller ikke beskrives på mikrostrukturelt plan, idet en mikrostruktur forudsætter eksistensen af en makrostruktur. (Makro- og mikrostruktur forstås her udelukkende som metaleksikografiske begreber. Jeg ser altså bort fra, hvordan homonyme termer bruges inden for tekstlingvistikken.) I min foreløbige analyse arbejder jeg med flg. grundtyper af megastruktur:

almindelig bog  
 (én) makrostruktur  
 flere makrostrukturer  
 makrostruktur, registerindgang  
 makrostruktur, registerindgang og saglig liste  
 makrostruktur, registerindgang og sproglig liste  
 makrostruktur, saglig liste  
 makrostruktur, sproglig og saglig liste  
 makrostruktur, sproglig liste  
 (= strukturel ordbogstypologi, 1. parameter)

På det makrostrukturelle plan udskilles typerne ikke efter komponenternes art og antal, men efter forskellige former for **organisering** af lemmerne. Et karakteristisk træk ved lemmerne er netop, at det er en komponent, som er rekursiv. Dette træk er måske det bedste bud på en strukturel invariabel i ordbogsbegrebet. Ingen makrostruktur er kendetegnende for bøger med "almindelig bog-megastruktur". Jeg har fundet flg. makrostrukturelle typer:

alfabetisk  
 alfabetisk nummereret  
 alfabetisk samt systematisk  
 ingen  
 numerisk  
 systematisk/billed  
 systematisk/tematisk  
 (= strukturel ordbogstypologi, 2. parameter)

På mikroplanet har jeg, ligesom ved megastrukturen, valgt at se på arten og antallet af mikrostrukturens/ordbogsartiklens komponenter. Man kan derfor med rette hævde, at jeg ikke - hverken i den mega- eller den mikrostrukturelle beskrivelse - beskæftiger mig med strukturen i snævrere forstand. Man kunne også sige det på en anden måde: den mega- og mikrostrukturelle beskrivelse fokuserer på træk ved **selektionen** af forskellige parametre (den såkaldte "indre selektion"), mens den makrostrukturelle ser på træk ved **præsentationen**. (Hvad angår den "ydre selektion", altså valget af lemmer til ordbogen, er der her tale om et anliggende for den materielle ordbogstypologi, jf. 3.1.). "Ingen" mikrostruktur er, ligesom "ingen" makrostruktur karakteristisk for bøger med "almindelig bog"-megastruktur. Jeg arbejder med følgende mikrostrukturelle træk:

definition  
 definition og forklaring  
 definition/synonym og forklaring  
 forklaring  
 ingen  
 synonym  
 synonym og definition  
 synonym og oplysninger til tekstproduktion  
 synonym og oplysninger til tekstreception  
 synonym, definition og oplysninger til tekstproduktion  
 synonym, definition og oplysninger til tekstreception



synonym, oplysninger til tekstproduktion og oplysninger til tekstreception  
 term, definition  
 ækvivalent  
 ækvivalent og definition  
 ækvivalent og oplysninger til tekstproduktion  
 ækvivalent og oplysninger til tekstreception  
 ækvivalent, definition og forklaring  
 ækvivalent, definition og oplysninger til tekstproduktion  
 ækvivalent, definition og oplysninger til tekstreception  
 ækvivalent, definition, oplysninger til tekstreception og oplysninger til tekstproduktion  
 ækvivalent, paradigmatisk leksik  
 ækvivalent, oplysninger til tekstreception og oplysninger til tekstproduktion  
 (= strukturel ordbogstypologi, 3. parameter)

### 3.3 Funktionel typologi

Den tredje og sidste typologiske dimension er den, som ser på hvordan ordbogen bruges - eller **kan** bruges - samt af hvem. Ud over disse to funktionelle parametre, brugen og brugeren, vil jeg nævne en tredje, som man gerne glemmer, nemlig ordbogens afsender. Jeg tror, at man i de kommende år vil se en øget opmærksomhed om afsenderen. Tænk blot på den omstændighed, at man i stadig større omfang omtaler ordbogens ophavsmand som "forfatter" og ikke blot som "sammenstillere". Der vil utvivlsomt i de kommende år, i takt med opmærksomheden på begrebet "brugervenlighed", blive stillet større krav til leksikografens evne til at indgå i dialog med brugeren. Herved forstår jeg øgede krav til, at leksikografen forudser i hvilke situationer, brugeren har behov for at slå op i en ordbog.

Med betegnelsen "afsender" hentydes ikke til den enkelte leksikograf, men til den, som har taget initiativet til ordbogen. Initiativtageren **kan** naturligvis være identisk med leksikografen, men er det kun sjældent. Afsenderen er altså snarere det forlag, den offentlige institution eller den private virksomhed, som har afdækket et behov, som det forsøger at dække med ordbogen. Afsenderen kan være delvist sammenfaldende med modtageren, nemlig når fx et firma for at tilfredsstille egne behov sætter et ordbogsprojekt igang. Særlig inden for bi- og polylingval fagleksikografi er det udbredt, at en international institution står bag udgivelsen af en ordbog. Målet er da ofte at fremme et fagligt samarbejde på tværs af landegrænser. I min gennemgang af de 67 fagordbøger har jeg fundet det rimeligt at arbejde med nedenstående typer af afsenderholdning:

belyse  
 diskutere  
 evaluere  
 fremme samarbejde  
 normere/standardisere  
 problematisere  
 reklamere  
 rådgive (ved oversættelse)  
 (= funktionel ordbogstypologi, 1. parameter)

Modtageren, eller rettere: den intenderede modtager af ordbogsteksten kan placeres efter sit faglige og sit sproglige niveau. Nogle af de nedenstående kategorier refererer til ét af disse parametre, andre til begge. Således er fx "ekspert" og "lægmand" udtryk for, at der hos modtageren forventes en større eller mindre grad af faglig viden, mens fx en "professionel oversætter" er en modtager med en stor sproglig og fag(sprog)lig viden. Følgende modtagerkategorier synes relevante for fagordbøgerne:

ekspert  
 korrespondent  
 lægmand  
 lægmand/fagmand  
 professionel oversætter  
 professionel oversætter/fagmand  
 professionel oversætter/studerende oversætter  
 sprogligt interesseret fagmand  
 tekniker  
 (= funktionel ordbogstypologi, 2. parameter)

Som det tredje funktionelle kriterium har jeg valgt at se på, hvilke situationer fagordbøgerne kan anvendes i. Forskningen i de situationer, som ordbøger - for slet ikke at tale om fagordbøger - kan bruges til, er endnu kun i sin vorden. De brugssituationer, eller "ordbogsanvendelsesformål", som anføres nedenfor, er derfor i allerhøjeste grad tentative. En videnskabelig redegørelse for ordbogsbrugssituationerne forudsætter, som Wiegand (1987:185) siger, at den enkelte situation opløses i sine komponenter, dvs. delsituationer. Af listen over situationer fremgår, at fagordbøgerne generelt kan bruges til meget forskellige ting. Den største forskel er den mellem ordbøger til læsning (bla. indlæring af sprog og/eller fag) og ordbøger til opslag, herunder oversættelse, tekstforståelse og forskellige former for fagligt arbejde.

baggrundsviden  
 baggrundsviden og fagindlæring  
 fagindlæring  
 fagindlæring og sprogindlæring  
 klassificering  
 medtænkning  
 oversættelse L1-L2  
 oversættelse L2-L1  
 oversættelse L2-L1, L1-L2  
 oversættelse L2-L1, L1-L2 og tekstreception  
 samarbejde  
 sprogindlæring  
 søgning  
 søgning og klassificering  
 tekstproduktion  
 tekstreception  
 tekstreception og fagindlæring  
 tekstreception og oversættelse  
 (= funktionel ordbogstypologi, 3. parameter)

Jeg understreger, at alle disse funktionelle klasser er opstillet på grundlag af en gennemgang af et tilfældigt, begrænset udvalg af fagordbøger. Det skal også understreges, at funktionelle træk i høj grad må udledes ved interpretation - i modsætning til de fænomenologiske (materielle og strukturelle) træk, som ifølge sagens natur er mere umiddelbart gennemskuelige. Ganske vist fremgår det undertiden af forord og brugervejledninger, hvad ordbogen skulle kunne bruges til. Det kan man så tage for pålydende - og så underkaste kritik senere. I andre tilfælde siges der ikke noget om funktionen, og hvis man ønsker at få de funktionelle træk med, er man således tvunget til at foretage sin egen fortolkning.

#### 4. Typologiske korrelationer - nogle konklusioner

Det interessante er nu at undersøge, hvilke korrelationer af materielt, strukturelt og funktionelt typologiske træk der er karakteristiske for de undersøgte fagordbøger.

De 67 undersøgte fagordbøger fordelte sig nogenlunde ligeligt på 26 monolingvale, 22 bilingvale og 19 polylingvale.

De største typologiske forskelle finder man inden for de **monolingvale** fagordbøger. Her er den faglige spredning stor: alle UDK-hovedgrupper undtagen nr. 9 (se 2.1.), som heller ikke optræder i de to andre kategorier af ordbøger, er repræsenteret. Desuden er 10 af de 16 materielle typer (og undertyper) med blandt de 26 udvalgte. 5 af dem er begrebsdefinitionsordbøger (Be og BeSa), alle har alfabetisk makrostruktur og deres vigtigste funktionelle træk er, at de skal formidle baggrundsviden og fagindlæring for en lægmand, fx:

**labile Gener**, Gener, som er særligt tilbøjelige til at mutere.

**latent Befrugtningsvirkning** betegner et specielt Fænomen hos visse Organismer (Snegle), hvor Moderindividet er bestemmende for visse Egenskaber ved Afkommet (Sneglehusets Drejning) uanset Afkommets individuelt afvigende Genotype, som først viser sig paa næste Generation igen.

**lateral Chiasma**, sidestillet Chiasma mellem to Kromosomer, der forbinder det ene Kromosom med det andets Langside til en Y-formet Forgrening.

(Wanscher 1942)

Det er også kendetegnende, at det kun er blandt de monolingvale ordbøger, at man finder sagordbøgerne (Sa, SaBe, SaBeSp og SaSp). Strukturelt præges de af en fattig megastruktur, en alfabetisk makrostruktur og en mikrostruktur, hvor forklaringen er det konstante element. Det, som brugeren, der altid er lægmand, kan anvende en sådan ordbog til, er primært at forøge sin encyklopædiske baggrundsviden om det pågældende faglige sagsområde. Undertiden satser afsenderen ikke blot på at oplyse, men også på at problematisere et emne og dermed på en måde indlede en dialog med læseren.

#### *alliance*

Forbund, sammenslutning eller samarbejde om et bestemt mål. NATO er en militær-alliance, d.v.s. et forbund mellem lande med det formål at forsvare *kapitalis-men* i medlemslandene og deres *kolonier*. Man kan også slutte forbund, alliere sig, med forretningsmæssige formål. Se *kartel*, *trust*, *monopol*.

*alternativ*

En anden mulighed, en anden løsning. *Socialismen* er det alternativ, som vil afløse *kapitalismen*. *Kollektiv* trafik er et alternativ til *privatbilismen*. Alternativ tænkning: tænkning i andre baner end de sædvanlige.

(Jesper 1974)

De 22 **bilingvale** befinder sig alle inden for hovedgrupperne 6 (anvendt videnskab, teknik) og 3 (samfundsvidenskaberne). 10 af dem er rene sprogordbøger (Sp), dvs. man finder kun ækvivalenter og andre rent sproglige elementer i deres mikrostruktur. Det kan forekomme overraskende, selv om de fleste henvender sig til professionelle oversættere, der således åbenbart formodes at have, eller have anden adgang til, oplysninger om den saglige og begrebsmæssige baggrund. Fordelingen af funktionerne L1>L2-oversættelse (aktiv) og L2>L1-oversættelse (passiv) er nogenlunde ligelig, og specielt må da den første formodes at stille krav til oversætterens encyklopædiske viden.

**forfaldsdag** Fälligkeitstag m.; Verfallstag m.; - 6 måneder efter ~ - 6 Monate (m.) seit dem Fälligkeitstag.

**forfaldstid** Fälligkeit f.; en veksels ~ - Verfallzeit (f.) eines Wechsels (m.); betaling før ~ - vorzeitige Zahlung.

(Gubba 1975)

Resten af de bilingvale ordbøger lader begrebsdefinitioner og saglig viden støtte de sproglige informationer.

Endelig gælder det for næsten samtlige **polylingvale** ordbøger, at de primært giver sproglig information, evt. kombineret med begrebsdefinitioner og saglig viden. Deres makrostruktur består typisk af en hovedliste, hvor ordene er opført efter et af de indgående sprog. Fra de andre sprog er der adgang til hovedlisten via et register, hvor man bliver henvist enten til et løbenummer eller til et lemma i det sprog, som hovedlisten er arrangeret efter. Nedenstående eksempel adskiller sig lidt fra mængden, idet makrostrukturen ikke er alfabetisk, men nummerisk. (Den nummeriske rækkefølge synes i det konkrete tilfælde ikke at afspejle et bestemt system, selvom der er tale om en vis indholdsmæssig progression.) Fra samtlige sprog er der adgang til hovedlisten via alfabetiske registre.

## 72. FUGEAFSTAND.

Afstand mellem to på hverandre følgende fuger.

## FUGENABSTAND.

Breite von einer Fuge zu der nächsten in einer Pflasterung oder in einem Betonlag.

## JOINT SPACING.

The predetermined distance between successive joints in a surfacing.

## SEPARACIÓN (DE LAS JUNTAS).

Distancia que existe entre dos juntas sucesivas.

## ÉCARTEMENT (ENTRE LES JOINTS).

Espace ménagé entre deux joint.

## SEPARAZIONE (DEI GIUNTI).

La distanza predeterminata fra due successivi giunti in una pavimentazione.

## 73. LÆNGDEFUGE, RÆKKEAFSTAND.

(Vejteknisk ordbog 1951: [makrostruktur])

627. Fuge.

72. Fugenabstand

186. Fugenbrett.

(Vejteknisk ordbog 1951: [tysk registerindgang])

## 5. Afslutning

Jeg har i denne artikel sigtet på at give nogle empirisk baserede informationer om danske fagordbøger. Omfanget af materialet gjorde, at jeg ikke her også kunne gå den vej, som jeg har beskrevet i en anden artikel (Mikkelsen 1991), hvor der tages udgangspunkt i funktionen, dvs fx oversættelse. Ud fra en kortlægning af de centrale faser i en sådan brugssituation kombineret med de øvrige funktionelle parametre, kan man så foretage en kritisk evaluering af ordbøgernes evne til at honorere de krav, som stilles til dem af brugeren. Generelt vil det være særdeles interessant at foretage en sammenligning mellem de materielle, strukturelle og funktionelle træk, som ordbogsudgiveren, direkte eller indirekte, hævder at ordbogen har - og de træk af samme karakter, som fremkommer ved den kritiske analyse. Kan man føle sig sikker på, at fagordbogen indeholder de rigtige fagord? - Er der så mange opslagsord/lemmer, som reklamerne siger? - Kan en fagordbog til oversættelse dække alle oversætterens problemer? For at kunne besvare disse spørgsmål og mange flere må der udvikles nogle værktøjer, som kan bruges i en konstruktiv kritik. Den historiske statusopgørelse over hidtidige fagordbøger er det første skridt i denne retning.

## Noter

1. Der er selvfølgelig også mulighed for at fokusere på andre dele af universet, fx på et område, som er fælles for en række faglige universer uden at indgå i det almene univers. Fx koncentrerer polytekniske ordbøger sig om et sådant udsnit. For en nærmere redegørelse for distinktionen mellem subsprog (=helsprog), særsprog og fællessprog henvises til Mikkelsen 1990.

## Primærlitteratur

Gubba, Wilhelm (1975): *Juridisk ordbog. Dansk-tysk*, Mårslet: W. Gubbas Forlag.

Jesper, Bente (1974): *Det lille røde leksikon*, København: Hans Reitzel.

*Vejteknisk Ordbog i 6 Sprog* (1951), Den Permanente Forening for Vejkongresser, Paris.

Wanscher, Johan Henrik (1942): *Arvelighedsordbog*, København: G.E.C. Gads Forlag.

**Sekundærlitteratur**

- Hausmann, Franz Josef (1989): Wörterbuchtypologie. I: Hausmann, Franz Josef / Oskar Reichmann / Herbert Ernst Wiegand / Ladislav Zgusta (eds.): *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography, First Volume*. Berlin, New York: Walter de Gruyter, 102-111 (= *Handbücher zur Sprach- und Kommunikationswissenschaft, Band 5.1*).
- Hausmann, Franz Josef / Oskar Reichmann / Herbert Ernst Wiegand / Ladislav Zgusta (eds.) (1989): *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography, First Volume*. Berlin, New York: Walter de Gruyter (= *Handbücher zur Sprach- und Kommunikationswissenschaft, Band 5.1*).
- Hupka, Werner (1989): Das enzyklopädische Wörterbuch. I: Hausmann, Franz Josef / Oskar Reichmann / Herbert Ernst Wiegand / Ladislav Zgusta (eds.): *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography, First Volume*. Berlin, New York: Walter de Gruyter, 988-999 (= *Handbücher zur Sprach- und Kommunikationswissenschaft, Band 5.1*).
- Mikkelsen, Hans Kristian (1990): Fagsprog og almensprog: særsprog og fællessprog. I: LSP *Nyhedsbrev* 2, 1-6.
- Mikkelsen, Hans Kristian (1991): Towards the ideal special language translation dictionary. I: *Hermes* 6, 91-109.
- Spang-Hanssen, Henning (1983): Fagsprog og almensprog. I: *Skrifter om anvendt og matematisk lingvistik* 10. Københavns Universitets Institut for anvendt og matematisk lingvistik, 85-100.
- Wiegand, Herbert Ernst (1987): Zur handlungstheoretischen Grundlegung der Wörterbuchbenutzungsforschung. I: *Lexicographica* 3, 178-227.

Sandro Nielsen

## Forholdet mellem alordbøger og enkeltfagsordbøger

For ikke-eksperter er den tosprogede fagalordbog af væsentlig betydning ved fx oversættelse af fremmedsproglige fagtekster og oversættelse af fagtekster til et fremmedsprog. I en sådan ordbog findes sproglige og encyklopædiske oplysninger om lemmata og ækvivalenter. Derfor bør man overveje at inkorporere kollokationer, eksempelsætninger, synonymi- og antonymiangivelser, samt illustrationer. Da sådanne oplysninger optager megen plads, bør man udarbejde tosprogede fagordbøger inden for snævre fagområder. Sådanne ordbøger kan give en tilbundsående behandling af sproget inden for de pågældende fagområder.

### 1. Indledning

Et af de væsentligste aspekter ved den tosproglige fagleksikografi er, at de fleste brugere af tosprogede fagordbøger kan betegnes som lægfolk, eller i bedste fald som halveksperter. En vigtig årsag til dette er, at der findes mange gode uddannelser inden for henholdsvis sprog og fag, men der findes ingen uddannelser, som kombinerer de to ting således, at de færdiguddannede er lige gode til både sprog og fag. Derfor spiller dette forhold en stor rolle inden for den tosproglige fagleksikografi. Dette skyldes ikke mindst det forhold, at sprogligt kyndige ordbogsbrugere generelt ikke besidder nogen væsentlig faglig viden om et eller flere fag inden for hvilke(t) de er beskæftiget med interkulturel kommunikation. Det er i denne forbindelse værd at bemærke, at de sprogligt kyndige ordbogsbrugere ofte hverken besidder faglig kompetence på modersmålet eller på fremmedsproget. En konsekvens heraf er, at en god tosproget fagordbog skal kunne hjælpe med relevante oplysninger om både sprog og fag inden for begge sprogområder. Hvordan en sådan tosproget fagordbog kan opbygges og hvilke typer af oplysninger den generelt skal indeholde, vil blive diskuteret i det følgende.

### 2. Typologi (1. del)

Som nævnt ovenfor er den ideelle tosprogede fagordbog en ordbog, der giver brugerne oplysninger om sprog såvel som fag. Inden for ordbogstypologien findes en type ordbog, der er specielt kendetegnet ved at indeholde oplysninger om både sprog og fag, og dette er **alordbogen**, også kaldet **albogen**. Senest har denne særlige type ordbog fået sin renæssance ved at blive fremhævet i Wiegand (1988) som værende den optimale ordbogstype inden for den fagsproglige leksikografi. Alordbogen, eller som den betegnes på tysk *Allbuch* (jf. Hausmann 1985 og Wiegand 1988) er en slags hybrid inden for ordbogsverdenen. Alordbogens hybride karakter er grundlæggende, at dens genuine formål er at give ordbogsbrugeren informationer om både sproglige og encyklopædiske aspekter vedrørende de lemmatiserede leksikalske enheder. I denne forbindelse er det vigtigt at holde sig for øje, at denne beskrivelse af alordbogen gælder for den ensprogede alordbog. Den tosprogede alordbog indeholder derimod sproglige og encyklopædiske oplysninger om både lemmata og ækvivalenter.

Det kan måske diskuteres, hvorvidt en ordbog, der indeholder sproglige og encyklopædiske oplysninger er en speciel ordbogstype, som ikke allerede er beskrevet i den metaleksikografiske litteratur. Det er nemlig muligt at argumentere, at en ordbog, der indeholder både sproglige og encyklopædiske oplysninger er en **encyklopædisk ordbog** (*encyclopedic dictionary, enzyklopädisches Wörterbuch*) eller måske endog en **encyklopædi**. Ifølge Zgusta (1971) er forskellen mellem sprogordbøger (*linguistic dictionaries, Sprachwörterbücher*) og encyklopædiske ordbøger som følger:

In the first place, we must differentiate encyclopedic dictionaries from linguistic ones. The latter are primarily concerned with the language, i.e. with the lexical units of language and their linguistic properties;[...] In contradistinction to this, the encyclopedic dictionaries (the biggest and most general of which are frequently called encyclopedias) are primarily concerned with the denotata of the lexical units (words): They give information about the extralinguistic world, physical or non-physical, and they are only arranged in the order of the words (lexical units) by which the segments of this extralinguistic world are referred to when spoken about. (Zgusta 1971:198)

Hvis man tager hensyn til kvalifikationen *primarily* i Zgustas karakteristik, synes det at være en rimelig betragtning, at en encyklopædisk ordbog indeholder to generelle artikeltyper. Den første artikeltype udgøres af de artikler, som indeholder typiske encyklopædiske oplysninger, fx vedrørende *proprier* (encyklopædiske artikler). Den anden artikeltype udgøres af de artikler, som indeholder typiske lingvistiske oplysninger, fx vedrørende fleksions- og derivationsmorfologi (lingvistiske artikler). LMED<sup>1</sup> er et eksempel på en encyklopædisk ordbog, hvis artikler i stort omfang lader sig opdele i disse to generelle kategorier. Dette er selvfølgelig en meget generel karakteristik og skal kun opfattes som sådan. Hvis man sammenligner med artiklerne i en alordbog, viser der sig imidlertid en forskel på artiklernes generelle indhold af informationer. I en typisk alordbogsartikel finder man nemlig både sproglige og encyklopædiske oplysninger i alle de tilfælde, i hvilke det er muligt og relevant at give ordbogsbrugeren sproglige og encyklopædiske oplysninger vedrørende de lemmatiserede leksikalske enheder (hybride artikler). I en tosproget alordbog er der også sproglige og encyklopædiske oplysninger vedrørende de til lemmaerne adresserede ækvivalenter.

Dette skal dog ikke opfattes som om der er et klart afgrænset skel mellem encyklopædiske ordbøger og sprogordbøger. Endvidere er der heller ikke noget skarpt skel mellem encyklopædiske ordbøger og alordbøger når man ser på de typer af oplysninger, som disse to ordbogstyper indeholder. At det kan være endog meget vanskeligt at skelne mellem, hvad der er sproglige oplysninger, og hvad der er encyklopædiske oplysninger i en fagordbogs artikler anskueliggøres gennem følgende udsagn:

Regardless of this multiplicity of dictionary types, one point cannot be stressed enough: a language's vocabulary reflects its speakers' knowledge of the world in which it is used. Any strict separation of linguistic-lexical and extralinguistic-factual information is very difficult, if not impossible. (Hartmann 1983:7)

<sup>1</sup> Se side 286 for forklaring på forkortning af primærlitteraturen



Afslutningsvis kan det nævnes, at en typisk alordbog adskiller sig fra en encyklopædisk ordbog, og derved også fra en encyklopædi, på to punkter. For det første, indikerer substantivet *encyklopædi* og adjektivet *encyklopædisk*, at et opslagsværk som fx en encyklopædi behandler et stort antal særskilte fagområder (LMED skelner således mellem 58 fagområder). Dette kan begrundes i ordenes græske oprindelse, idet den oprindelige græske betydning - som tilsyneladende stadig er den gængse betydning - af ordet *encyklopædi* indikerer en beskrivelse af alle videnskabernes enkeltdiscipliner, eller fagområder (jf. GFO). Denne forskel understreges endvidere af, at de 48 encyklopædiske ordbøger og encyklopædier nævnt i Hupka (1989) tilsyneladende alle behandler et stort antal selvstændige fagområder. Til forskel herfra kan en alordbog nøjes med at behandle et enkelt fagområde (se også pkt. 4.3 og pkt. 5.).

For det andet, er et af de karakteristiske træk ved en encyklopædi og en encyklopædisk ordbog ifølge Hupka (1989), at disse to typer af opslagsværker indeholder et stort antal proprier:

(b) Enzyklopädie: Vermittlung von Sachwissen (enzyklopädische Information), in der Regel Eigennamenartikel.

(c) Enzyklopädisches Wörterbuch: Vermittlung von Fachwissen (enzyklopädische Information), Definition/Bedeutungsbeschreibung zur Vermittlung eines durchschnittlichen Allgemeinwissens, Angaben zur Sprache, meist Eigennamenartikel. (Hupka 1989:992)

I modsætning hertil er det ikke noget karakteristisk træk ved en alordbog, at den indeholder proprier blandt sine lemmata. Det er praktisk og teoretisk muligt - og slet ikke usandsynligt - at en alordbog overhovedet ikke indeholder proprier blandt sine lemmata. Dette betyder bl.a., at to af de væsentlige karakteristika, som kendetegner en encyklopædisk ordbog ikke er karakteristiske træk for en alordbog. Det skal dog bemærkes, at der her ikke sættes lighedstegn mellem en encyklopædisk ordbog og en alordbog, men at de af praktiske årsager holdes adskilt som to forskellige ordbogstyper.

Ovenfor er alordbogen blevet placeret i en simpel ordbogstypologi, og i det følgende vil de enkelte former for oplysninger, der er indeholdt i den typiske alordbog blive diskuteret.

### 3. Alordbogens oplysningstyper

Som nævnt ovenfor er der mere end én type oplysninger i den typiske alordbog. I forbindelse med den tosprogede alordbog betyder dette bl.a., at ordbogen skal indeholde mere end blot lemmata og ækvivalenter (sml. IFF). Det bør bemærkes, at en ordbog skal være udarbejdet til at opfylde et bestemt formål (fx oversættelse) og for en bestemt målgruppe (fx translatør-studerende) for i praksis at kunne leve op til dette krav. For at begrænse diskussionen vil der i det følgende blive taget hensyn til en målgruppe, der udgøres af den sprogligt kyndige lægmand/-kvinde. Som det vil fremgå nedenfor har dette betydning for de oplysninger, der generelt bør inkorporeres i en tosproget fagalordbog.

Den tosprogede fagalordbog er en hybrid ordbogstype, idet den indeholder både sproglige og encyklopædiske oplysninger. I et forsøg på at anskueliggøre det nærmere indhold af disse to oplysningstyper, henvises til Wiegand (1988) hvor alordbogens informationstyper er gjort til genstand for en nærmere undersøgelse. Dette betyder endvidere, at den af Wiegand fremsatte definition af **sproglige oplysninger** i alordbogen vil blive overtaget:

Standardisierte linguistische Angaben sind solche funktionale lexikographische Textsegmente, deren genuiner Zweck darin besteht, daß der Benutzer aus den angabeinternen Daten Informationen über sprachliche Eigenschaften des Lemmazeichens (oder eines anderen Textsegmentes) gewinnen kann. (Wiegand 1988:766)

*Funktionale lexikographische Angaben* kan fx være: oplysninger vedrørende fonologi, fleksionsmorfologi, ordklasse, orddeling og valens (jf Wiegand 1988:766 og Scholfield 1980:102). Det fremgår af denne langt fra udtømmende opremsning, at sproglige oplysninger er adresserede til formen og indholdet af et lingvistisk tegn repræsenterende en leksikalsk enhed. Med hensyn til hvilke typer sproglige oplysninger, der kan inkorporeres i fagordbøger, henvises endvidere til Mugdan (1989) og Scholfield (1980).

Med hensyn til de **encyklopædiske oplysninger**, som ideelt set bør være indeholdt i en alordbog, hvad enten det er i en hybridartikel eller et andet sted i ordbogen, er det fundet hensigtsmæssigt at overtage den i Wiegand (1988) fremsatte definition:

Denn parallel zu den linguistischen können die standardisierten enzyklopädischen Angaben als solche funktionale lexikographische Textsegmente bestimmt werden, deren genuiner Zweck darin besteht, daß der potentielle Benutzer aus den angabeinternen Daten Informationen über Eigenschaften eines Gegenstandes sowie solche zu Handlungskontexten, in der dieser seine Zwecke erfüllt, gewinnen kann, wobei es sich um denjenigen gegenstand handelt, auf den man mit denjenigen Lemmazeichen usuell Bezug nimmt, das im gleichen Wörterbuchartikel als Lemma fungiert, in dem die enzyklopädische Angabe steht. (Wiegand 1988:766)

Som det fremgår af denne definition er encyklopædiske oplysninger ikke relaterede til det sproglige tegns udtryksside eller indholdsside. Derimod er *encyklopædiske Angaben* relaterede til faktuelle egenskaber vedrørende leksikalske enheder. Sådanne oplysninger er især relevante for ordbogsbrugere, der ikke besidder nogen fagviden eller kun besidder begrænset fagviden vedrørende de i ordbogen inkorporerede lemmata, ækvivalenter og/eller det fagområde, som behandles i ordbogen.

Ovenfor er i generelle vendinger beskrevet hvad der forstås ved sproglige og encyklopædiske oplysninger, men det er relevant at se nærmere på disse to typer af oplysninger. For yderligere at illustrere dette aspekt, kan de sproglige oplysninger i en alordbogsartikel anskueliggøres gennem følgende citat:

X ist eine sprachliche Angabe zur Eigenschaft E des Lemmazeichens Y.  
(Wiegand 1988:764)

I denne definition repræsenterer *X* det funktionelle leksikografiske tekstsegment, *E* repræsenterer en specifik egenskab ved lemmategnet, og *Y* repræsenterer et givet lemmategn. Lemmategnet er det sproglige tegn, der repræsenteres grafisk af et lemma i en ordbogsartikel. Det ovenstående udsagn er dog kun sandt for så vidt angår ensprogede alordbøger, hvilket Wiegand også retteligt gør opmærksom på. Årsageren til dette er, at artiklerne i en tosproget alordbog udover lemmata også indeholder ækvivalenter, dvs. både udgangssproglige og målsproglige elementer. For at få anskueliggjort dette, er det nødvendigt at udvide ovenstående definition således, at den kommer til at indeholde følgende elementer:

$X$  er en sproglig angivelse, der betegner egenskaben  $E$  ved lemmategnet  $Y$ , og  $X_1$  er en sproglig angivelse, der betegner egenskaben  $E_1$  ved ækvivalenttegnet  $Z$ .

I denne udvidede definition repræsenterer  $X_1$  det funktionelle leksikografiske tekstsegment,  $E_1$  repræsenterer en specifik sproglig egenskab ved ækvivalenttegnet, og  $Z$  repræsenterer et givet ækvivalenttegn. Et ækvivalenttegn er det sproglige tegn, der repræsenteres grafisk af en ækvivalent i en ordbogsartikel. Med andre ord: i en tosproget alordbog er der sproglige oplysninger om både lemmæerne (lemmategnene) og ækvivalenterne (ækvivalenttegnene). På linie med den ovenstående definition af sproglige oplysninger i en alordbog, er det muligt at beskrive de relevante encyklopædiske oplysninger i en tosproget alordbog på følgende måde:

$X$  er en encyklopædisk angivelse, der betegner egenskaben  $E$  ved lemmategnet  $Y$ , og  $X_1$  er en encyklopædisk angivelse, der betegner egenskaben  $E_1$  ved ækvivalenttegnet  $Z$ .

I denne definition repræsenterer de variable størrelser  $X$  og  $X_1$  hver sit funktionelle leksikografiske tekstsegment,  $E$  og  $E_1$  repræsenterer hver sin specifikke encyklopædiske egenskab ved henholdsvis lemmategnet og ækvivalenttegnet,  $Y$  repræsenterer et givet lemmategn og  $Z$  repræsenterer et givet ækvivalenttegn. På baggrund af de to sidst anførte definitioner er det muligt at illustrere oplysningerne i en tosproget alordbogsartikel ved hjælp af følgende skematiske modelartikel:

Art. 1:        lemma; sproglige oplysninger vdr. lemma; encyklopædiske oplysninger vdr. lemma; ækvivalent; sproglige oplysninger vdr. ækvivalent; encyklopædiske oplysninger vdr. ækvivalent.

I denne forbindelse bør det bemærkes, at ovenstående skematiske modelartikel (Art. 1) på ingen måde repræsenterer en udtømmende opremsning af mikrostrukturelle enheder i en given alordbogsartikel. Dette skyldes bl.a., at en tosproget alordbog, hvad enten dens genuine formål er tekstreception eller tekstproduktion, ideelt set bør indeholde flere elementer end de ovenfor nævnte (jf. fx Moulin 1983:149-151). Med udgangspunkt i Bergenholtz/Poulsen (1989) kan man udvide alordbogsartiklens mikrostruktur med følgende elementer (jf. Bergenholtz/Poulsen 1989:112-113):

1. oplysninger om kollokationer
2. eksempelsætninger
3. synonymiangivelser
4. antonymiangivelser

Heraf følger, at en alordbogsartikel kan indeholde et betydeligt antal forskellige oplysningstyper. Udover sproglige og encyklopædiske oplysninger vedrørende lemma og ækvivalent, må man tage højde for også at skulle inkorporere en, flere eller alle af disse typer oplysninger i en hybridartikel i en tosproget alordbog. Et af de vægtigste argumenter for inkorporeringen af disse typer af oplysninger er, at den fagligt ukyndige bruger ikke kan forventes at være bekendt med de inden for et eller flere fag forekomne ord og vendinger på

både modersmålet (L1) og fremmedsproget (L2). Dette gælder i særdeleshed faste ordforbindelser (kollokationer). For det andet vil eksempelsætninger i mange tilfælde med fordel kunne anvendes til at illustrere brugen af ækvivalenter. Endvidere vil oplysninger om synonymi- og antonymirelationer kunne hjælpe den fagligt ukyndige bruger til at forstå nogle af de sammenhænge, der findes inden for et vidensområdes fagsystematik. Udover disse fire typer af oplysninger vil det i mange tilfælde også være relevant at inkludere krydshenvisninger til andre artikler i en ordbogs ordliste (*word list*, *Wörterverzeichnis*) eller til andre dele af ordbogen.

Sproglige og encyklopædiske oplysninger kan i flere tilfælde med fordel placeres i en ordlisteekstern, men samtidig ordbogsintern position. Hvad angår de sproglige oplysninger kan man i tosprogede fagalordbøger beregnet til oversættelse fra L1 til L2 inkorporere en selvstændig ordbogsgrammatik (jf. AND p. 10 og Bergenholtz 1990:26f). I denne ordbogsgrammatik kan man således behandle de regelmæssige grammatiske regler og/eller de grammatiske regler, der gælder for et større antal lemmata og/eller ækvivalenter. Ideen bag dette princip er, at man således undgår at inkorporere den samme grammatiske oplysning i mange ordbogsartikler.

Hvad angår de encyklopædiske oplysninger kan man i tosprogede fagalordbøger inkorporere en **komparativ oversigt** over et fags systematik (jf. Bergenholtz/Dam/Henriksen 1990:135). I denne komparative fagsystematiske oversigt kan fagets grundlæggende regler for både L1 og L2 sprogområdet beskrives. Dette er specielt vigtigt ved kulturspecifikke fagområder som fx engelsk og dansk ret. Såfremt denne oversigt inddeles i særskilte, nummererede paragraffer (jf. AND), er det muligt at benytte et krydshenvisningssystem fra de enkelte ordbogsartikler til den fagsystematiske del. Dette betyder bl.a., at man undgår en unødigt belastning af de enkelte artiklers mikrostruktur således, at der i artiklerne er plads til andre relevante oplysninger.

#### 4. Typologi (2. del)

I forbindelse med både ensprogede og tosprogede ordbøger har der i tidens løb været et ikke ubetydeligt antal forsøg på at opstille en almenyldig klassifikation af ordbøger i forhold til forskellige kriterier (jf. fx Zgusta 1971, Malkiel 1975, Al-Kasimi 1977, Kromann/Riiber/Rosbach 1984 og Hausmann 1989). Det er imidlertid værd at bemærke, at ingen af disse klassifikationer eller typologier er opstillet med særlig henblik på fagsproglige ordbøger. I det følgende vil der blive taget udgangspunkt i en slags formålsbeskrivelse. Med dette udgangspunkt vil der nedenfor blive gjort forsøg på at identificere den eller de typer af ordbøger, der er bedst egnede til at opfylde de krav, som en tosproget fagalordbog stiller med hensyn til de ovenfor nævnte typer af informationer.

##### 4.1 Lemmabestand

I forbindelse med fagsprogsleksikografien kan man opdele ordbøgerne i to forskellige typer. Man skelner mellem maksimerende (el. ekstensive) ordbøger og minimerende (el. selektive) ordbøger (jf. Hausmann 1977:5f). Ved en **maksimerende** fagsproglig ordbog forstås en ordbog, der er beregnet til at indeholde så stor en del af et på forhånd defineret fagsprogligt ordforråd som muligt. Ved en **minimerende** fagsproglig ordbog forstås en ordbog, der er beregnet til kun at indeholde en bevidst begrænset del af et på forhånd defineret fagsprogligt

ordforråd. Et eksempel på en maksimerende fagsproglig ordbog kunne være en engelsk-dansk juridisk ordbog, der indeholder så mange af de i engelsk juridisk sprog forekomne ord og vendinger som overhovedet muligt. Et eksempel på en minimerende fagsproglig ordbog kunne være en engelsk-dansk juridisk ordbog, der indeholder de 100 mest frekvente fagord inden for engelsk juridisk sprog.

I tilknytning til denne dikotomi anser Hausmann en ordbog, der indeholder mere end 30.000 lemmata for at være en maksimerende ordbog og en ordbog, der indeholder mindre end 30.000 lemmata for at være en minimerende ordbog (Hausmann 1977:5f). Det er dog vigtigt at holde sig for øje, at dette grænsetal gælder for "almensproglige" ordbøger og ikke for fagsproglige ordbøger. Hvorvidt dette tal er det korrekte tal kan diskuteres, men én ting synes givet, nemlig at antallet af leksikalske enheder i et givet sprogs "almene" ordforråd er større end antallet af leksikalske enheder i et bestemt afgrænset fagområde inden for det samme sprog (jf. Nielsen 1989:60).

#### 4.2 Fagordbøgernes fagområde

Opdelingen af fagsproglige ordbøger som enten maksimerende eller minimerende er dog ikke tilstrækkelig til beskrivelse af forskellige typer af fagordbøger. Tosprogede — og også ensprogede og flersprogede — fagordbøger kan inddeles i to selvstændige ordbogstyper. Den første af disse ordbogstyper er **enkeltfagsordbogen** (*single-field dictionary*), hvorved forstås en fagsproglig ordbog, der er beregnet til at dække det ordforråd, som benyttes inden for ét selvstændigt fagområde, fx juridisk sprog (jf. Nielsen 1990:132).

Når man har med enkeltfagsordbøger at gøre, kan man skelne mellem yderligere to typer af fagordbøger, nemlig den såkaldte **hovedfagsordbog** (*general-field dictionary*) og den såkaldte **delfagsordbog** (*sub-field dictionary*). Ved en hovedfagsordbog forstås en enkeltfagsordbog, der er beregnet til at dække ordforrådet inden for et enkelt fagområde, fx en ordbog over juridisk sprog (jf. Nielsen 1990:132). En sådan fagordbogs formål er at beskrive det ordforråd, der benyttes inden for et sådant fagområde generelt. Dette betyder bl.a., at en juridisk ordbog skal dække alle de selvstændige underområder (delfag), som findes inden for fagområdet jura, og derfor skal den indeholde leksikalske enheder fra fx underområderne strafferet, erstatningsret, familieret, procesret og arveret.

Ved en delfagsordbog forstås en enkeltfagsordbog, der er beregnet til at dække ordforrådet inden for et underordnet fagområde, fx en aftaleretlig ordbog (jf. Nielsen 1990:132). En sådan ordbogs formål er at beskrive det ordforråd, der benyttes inden for et enkelt selvstændigt underområde af et overordnet fagområde. Hvis man tager den nævnte aftaleretlige ordbog som eksempel betyder dette, bl.a., at en aftaleretlig ordbog indeholder lemmata fra dette juridiske underområde. Her må man dog tænke sig, at en delfagsordbog også vil indeholde lemmatiserede leksikalske enheder fra det pågældende overordnede fagområde, da de enkelte underområder inden for et fag til en vis grad griber ind i hinanden. En følge heraf er, at en delfagsordbog også må forventes at indeholde et større eller mindre antal leksikalske enheder, hentet fra det overordnede fagområde.

Den anden hovedtype af fagsproglige ordbøger er **flerfagsordbogen** (*multi-field dictionary*), hvorved forstås en ordbog, der er beregnet til at dække to eller flere overordnede fagområder inden for det, som med et bredt udtryk kaldes fagsproglig kommunikation (jf. Nielsen 1990:133). En sådan ordbogs formål er at beskrive det ordforråd, der benyttes ved kommunikation inden for de valgte fagområder som en flerfagsordbog omfatter. Et eksempel på en flerfagsordbog er den "traditionelle" erhvervsfaglige ordbog (*business dictionary*). En

fagsproglig ordbog af denne type indeholder lemmata fra flere forskellige selvstændige hovedfagområder, som fx fagområderne finansiering, management, transport, bankvæsen og forsikringsvæsen (jf. LDBE). Både enkeltfagsordbøger og flerfagsordbøger kan være minimerende og maksimerende, men denne skelnen er hovedsageligt kun relevant for enkeltfagsordbøger (jf. Nielsen 1990:133).

### 4.3 Et eller flere fag?

Efter at have set, at fagsproglige ordbøger kan inddrages i ordbogstyper efter antallet af de fag som de dækker og efter det dækkede fagområdes størrelse, er det på sin plads at se på nogle af de enkelte ordbogstypers karakteristika. Som udgangspunkt for den følgende oversigt er valgt dikotomien flerfagsordbøger/enkeltfagsordbøger.

For det første kan fler- eller multifagsordbøger potentielt indeholde et større antal lemmata end enkeltfagsordbøger. Dette kan begrundes ved, at det samlede ordforråd inden for to eller flere hovedfagområder efter al sandsynlighed er større end ordforrådet inden for et enkelt fagområde. Heraf kan man drage den slutning, at flerfagsordbøger tenderer til at have en større lemmabestand end enkeltfagsordbøger. Et eksempel herpå er IFF med (angiveligt) 150.000 lemmata over for JOW med ca. 20.000 lemmata (jf. også Bergenholtz/Nielsen 1988:252).

For det andet har en flerfagsordbog et større anvendelsespotentielt end en enkeltfagsordbog. Da en flerfagsordbog dækker flere end ét fagområde er dens anvendelsesmuligheder større - i det mindste i teorien - selv om man må holde sig for øje, at lemmaselektionen sætter sin naturlige begrænsning på ordbogens anvendelsespotentielt. Da en enkeltfagsordbog pr. definition kun dækker et enkelt fagområde er dens anvendelsesmuligheder hovedsageligt begrænset til dette specifikke fagområde. Også i forbindelse med enkeltfagsordbøger sætter lemmaselektionen sin begrænsning på ordbogens anvendelsesmuligheder. Man kan sige, at en enkeltfagsordbogs anvendelsespotentielt er fagintern, mens en flerfagsordbogs anvendelsespotentielt er faggrænseoverskridende.

For det tredje giver flerfagsordbogen en mere overfladisk leksikografisk behandling af de enkelte fagområders ordforråd end tilfældet er med flertallet af enkeltfagsordbøger. Dette skyldes bl.a., at det overvejende flertal af flerfagsordbøger dækker et relativt stort antal fagområder, fx dækker LDBE 25 selvstændige fagområder (jf. p. xii).

I tilknytning til den kvalitative behandling af lemmata, bør det nævnes, at den forholdsvis overfladiske leksikografiske behandling af de enkelte fagområder i flerfagsordbøgerne resulterer i en kvantitativ ringe behandling af fx kollokationer, synonymi- og antonymirelationer. I det mindste inden for den tosprogede fagleksikografi er det almindeligt anerkendt, at behandlingen af kollokationer skal tillægges betydelig vægt (jf. fx Scholfield 1980:104 og Cowie 1983:140f). Ganske vist er behandlingen af kollokationer forskellig fra ordbog til ordbog, men der synes at være en generel tendens til, at flerfagsordbøger indeholder et relativt stort antal kollokationer fra et lille antal fagområder på bekostning af de øvrige fagområder, der behandles. Problemet er her, at såfremt en flerfagsordbog skal indeholde alle kollokationer inden for et fagområde, vil resultatet blive en voluminøs, uhåndterlig og brugeruvenlig fagsprogsordbog.

På denne baggrund voves den påstand, at det leksikografisk set er bedre at udarbejde enkeltfagsordbøger end flerfagsordbøger, da enkeltfagsordbøger er ideelle til at give en dybtgående og detaljeret leksikografisk behandling af et i øvrigt vanskeligt tilgængeligt fagområde. Endvidere vil det være betydeligt lettere at udarbejde en alordbog, der dækker et

enkelt fagområde end flere fagområder. Den primære årsag til dette er, at en sådan alordbog indeholder et stort antal af forskellige oplysningstyper. I forbindelse med tosprogede fagalordbøger er der omfattende oplysninger vedrørende både udgangssprog og målsprog, samt både lingvistik og fag. Alle disse oplysninger kan bedst gives i forbindelse med et enkelt fag i modsætning til et stort antal forskellige fag, og følgelig er den optimale alordbog snarere en enkeltfagsordbog end en flerfagsordbog (jf. også pkt. 5).

#### 4.4 Hoved- eller delfagområde?

Dikotomien hovedfagsordbog/delfagsordbog vil nu blive gjort til genstand for en kort diskussion. For det første har en hovedfagsordbog muligheden for at indeholde et større antal lemmata end en delfagsordbog. Dette kan bl.a. begrundes ved, at fx det generelle fagområde jura har et større ordforråd end fx underområdet aftaleret (sml. den ovenfor nævnte forskel mellem flerfagsordbøger og enkeltfagsordbøger). Endvidere har hovedfagsordbøger et større anvendelsespotentielt end delfagsordbøger, da de førstnævnte ordbøgers lemmabestand dækker et større fagområde.

For det andet kan der spores en markant tendens til, at delfagsordbøger giver ordbogsbrugeren en udførlig og detaljeret leksikografisk behandling af det pågældende delområdes ordforråd i modsætning til den mindre detaljerede behandling af et hovedfagområde, der findes i flertallet af de eksisterende hovedfagsordbøger. Denne udførlige leksikografiske behandling af et delfagområde kan, og bør, resultere i inkorporeringen af et stort antal kollokationer, da der er mere plads i en delfagsordbog end i en hovedfagsordbog til disse vigtige syntagmatiske relationer. Det synes derfor rimeligt at fremføre den påstand, at en delfagsordbog er særlig velegnet til at give en detaljeret leksikografisk behandling af de forskellige sproglige og faglige aspekter for et begrænset fagområde. Dette betyder, at den mest udførlige og detaljerede leksikografiske behandling af et fagområde, som fx jura, opnås ved udarbejdelsen af et større antal maksimerende delfagsordbøger, fx en aftaleretlig ordbog, en erstatningsretlig ordbog, en familieretlig ordbog, en strafferetlig ordbog, etc., som tilsammen dækker hele det juridiske fagområde.

For det tredje vil en delfagsordbog være bedre egnet som en fagsproglig alordbog end en hovedfagsordbog. Dette beror bl.a. på, at de omfattende sproglige og encyklopædiske oplysninger i en fagsproglig alordbog optager forholdsvis megen plads. En delfagsordbog i skikkelse af en fagsproglig alordbog vil derimod være af et mere håndterligt fysisk omfang, eventuelt noget i retning af en maksimerende hovedfagsordbog, der ikke er en alordbog. Følgelig voves den påstand, at en fagsproglig alordbog af praktiske årsager snarere bør være en (maksimerende) delfagsordbog end en (maksimerende) hovedfagsordbog. Dette er dog også afhængigt af det pågældende fagområdes ordforråd.

#### 5. Alordbogen som enkeltfagsordbog

Det fremgår af det foregående, at enkeltfagsordbogen er bedre egnet som en alordbog end tilfældet er med flerfagsordbogen. Dette gælder især i de tilfælde, i hvilke fagalordbogen er en maksimerende ordbog. Hvad angår tosprogede fagordbøger er det både L1 og L2 aspekter, der skal beskrives inden for et givet fagområde og heraf følger, at der generelt kræves mere plads til at inkorporere alle relevante oplysninger i en maksimerende tosproget fagalordbog end i en maksimerende ensproget fagalordbog.

Da forskellen mellem "almindelige" tosprogede fagordbøger og tosprogede fagalordbøger især er iøjnefaldende i forbindelse med ordbøgernes artikler, vil udgangspunktet for denne korte sammenligning være artikelinterne oplysninger. Dette betyder dog ikke, at artikeleksterne oplysninger er irrelevante i denne sammenhæng, men her henvises af praktiske grunde til den ovenfor anførte diskussion af ordlisteeksterne men ordbogsinterne oplysninger under pkt. 3. Til illustration af den grundlæggende forskel mellem artikelinterne oplysninger i "almindelige" tosprogede fagordbøger og tosprogede fagalordbøger er det fundet hensigtsmæssigt at se på følgende to ordbogsartikler fundet i to engelsk-danske fagordbøger:

Art. 2-3                      offer tilbud. (FEO)  
   offer tilbud (IFF)

Som det fremgår af ovennævnte to eksempler, er artiklerne informationsmæssigt identiske. I lemmapositionen (således som Wiegand 1983 forstår det) findes et lemma og i ækvivalentpositionen findes en ækvivalent. Endvidere indeholder disse to artikler to mikrostrukturelle elementer. I lemmapositionen findes en retskrivningsangivelse (således som Wiegand 1989 forstår det), og det samme er tilfældet med ækvivalentpositionen. Ingen af de to artikler indeholder yderligere oplysninger, der kunne have været til hjælp for en potentiel ordbogsbruger. Art. 2 og Art. 3 er typiske eksempler, der er velegnede til at illustrere en ordbogsartikel i en "almindelig" engelsk-dansk fagordbog.

I en tosproget fagalordbog udarbejdet med særlig henblik på den fagligt ukyndige ordbogsbruger, er der som ovenfor nævnt behov for et væsentligt større antal oplysninger i artiklerne end i de to artikeleksempler fra FEO og IFF. En væsentlig forskel på en artikel i en "almindelig" tosproget fagordbog og i en tosproget fagalordbog er bl.a. inkorporeringen af encyklopædiske oplysninger og kollokationer i sidstnævnte. Med udgangspunkt i Art. 2 og Art. 3 vedrørende lemmaet *offer*, vil der nedenfor blive gennemført en kort sammenligning mellem en eksemplarisk alordbogsartikel og en "almindelig" fagordbogsartikel. Oplysningerne i en tosproget alordbogsartikel kan illustreres ved hjælp af følgende imaginære artikel fra en ligeledes imaginær engelsk-dansk juridisk ordbog udarbejdet for danske brugere:

Art. 4    *offer* n [erklæring fra en person om, at han er villig til at give eller afstå fra at gøre noget, og som kan accepteres af modtageren - accepteres erklæringen foreligger en bindende aftale]

*tilbud* n <pl ~>

• man er i almindelighed ikke forpligtet til at besvare tilbud, der uopfordret tilsendes én

△ *invitation to make an offer* opfordring til at gøre tilbud; *an offer becomes effective when it reaches the offeree* et tilbud får virkning når det kommer frem til tilbudsmotageren; *offer implied by conduct* stiltiende tilbud; *an offer is made when it actually reaches the offeree* et tilbud er afgivet når det er kommet frem til modtageren; *the offer has lapsed* tilbudet er bortfaldet; *offer to the world at large* generelt tilbud; *the offer remains open against a reply by telegram* tilbudet er fast mod telegramsvar; *once rejection takes place the offer is terminated* når et tilbud afslås er det endeligt bortfaldet; *he rejected the offer* han afslog tilbudet; *he revoked the offer* han tilbagekaldte tilbudet; *unsolicited offer* uanmodet tilbud

⇒ invitation to treat



Hvad angår sproglige oplysninger adresseret til lemmaet er der i Art. 4 retsskrivningsangivelse, ordklassangivelse, fonetisk angivelse og ordklasseangivelse. Hvad angår sproglige oplysninger adresseret til ækvivalenten er der retsskrivningsangivelse, ordklasseangivelse og fleksionsangivelse. Herudover er der en betydningsangivelse med encyklopædisk indhold adresseret til lemmaet. Artiklen indeholder en autentisk eksempelsætning, der viser brugeren af ækvivalenten, bl.a. det forhold, at singularisformen og pluralisformen af ækvivalenten er identiske (nulfleksion). Herefter følger 11 kollokationer med dertil hørende ækvivalentangivelser. Sluttelig findes en krydshenvisning til et andet lemma i den imaginære ordbog. Denne artikel er et eksempel på hvilke oplysninger, der kan være relevante i en tosproget fagalordbog såfremt man skal kunne løse de fleste problemer ved hjælp af kun en enkelt ordbog. Det er dog vigtigt, at leksikograferne inden udarbejdelsen af ordbogen har klarlagt hvilke behov ordbogens målgruppe har i forbindelse med den pågældende fagordbog, således at alle nødvendige oplysninger medtages i ordbogen.

Sammenlignes Art. 2 og Art. 3 med den imaginære alordbogsartikel anført ovenfor fremgår det med al tydelighed, at informationsmængden er betydelig større i Art. 4 end i de to førstnævnte artikler. Det er selvfølgelig svært - om ikke umuligt - at angive det eksakte antal af oplysninger, som bør inkorporeres i en ordbogsartikel i den optimale tosprogede fagalordbog. Man må dog som hovedregel kunne gå ud fra, at der som minimum skal være grammatiske oplysninger til lemma og ækvivalent, encyklopædiske oplysninger til (mindst) lemmaet og en eksempelsætning, der viser brugen af ækvivalenten. Herudover kommer et ikke definerbart antal kollokationer. Hvorvidt der skal være yderligere oplysninger i alle standardiserede ordbogsartikler kan diskuteres, men som ovenfor nævnt, bør leksikograferne overveje inkorporeringen af synonymiangivelser og antonymiangivelser samt eventuelle krydshenvisninger og illustrationer.

## 6. Afslutning

Et betydeligt antal af de eksisterende fagordbøger, hvad enten det drejer sig om ensprogede eller tosprogede ordbøger, dækker de forskellige fagområder på utilstrækkelig vis. Især de tosprogede fagordbøger indeholder et utilstrækkeligt antal lemmata, hvilket resulterer i, at ordbogsbrugerne ofte må lede forgæves i ordbøgerne (dvs. minimerende ordbøger). Hertil kommer, at de tosprogede fagordbøger indeholder et utilstrækkeligt antal oplysninger vedrørende det eller de fag, ordbøgerne prætenderer at behandle. Den typiske engelsk-danske fagordbog indeholder kun lemmata og dertil adresserede ækvivalenter, men fx ingen fagsystematiske oplysninger, ingen grammatiske oplysninger og ingen ækvivalensoplysninger.

I en tosproget fagordbog, der skal kunne benyttes ved oversættelse til et fremmedsprog og i en ordbog, der skal kunne benyttes ved oversættelse af fremmedsproglige fagtekster til brugerens modersmål er sproglige og faglige oplysninger nødvendige. Endvidere er kollokationer og illustrationer i fagsproglige sammenhænge med til at fuldende beskrivelsen af sproget og systematikken inden for et givet fag. En ordbogstype, der kan opfylde i hvert fald en stor del af disse krav er en maksimerende fagalordbog i form af en delfagsordbog. En sådan ordbog vil sikre en udtømmende beskrivelse af det fagsprog, som benyttes inden for det pågældende delfagområde, idet dens muligheder for at inkorporere de nødvendige oplysninger generelt er større end ved en hovedfagsordbog. Under alle omstændigheder synes det at være utopi at forvente udarbejdelsen af en maksimerende flerfagsordbog som alordbog, hvorfor man bør koncentrere sig om udarbejdelsen af tosprogede fagalordbøger i form af enkeltfagsordbøger.

## 7. Litteratur

### 7.1 Primære kilder

- AND= Anderson, Ralph James Bushnan 1977: *Anglo-Scandinavian Law Dictionary of Legal Terms Used in Professional Practice*. Oslo [etc.]: Universitetsforlaget.
- COL= Collin, P.H. 1986: *English Law Dictionary*. Teddington: Peter Collin Publishing.
- GFO= Brüel, Sven 1975: *Gyldendals Fremmedordbog*. København: Gyldendal.
- FEO= Nielsen, Feodor 1975: *Anglo-amerikansk/dansk retsordbog*. København: Nyt Nordisk Forlag Arnold Busk.
- IFF= Høedt, Jørgen et al. 1988: *IFF fagordbog dansk/engelsk-engelsk/dansk, vols. 1-2*. København: IFF Ordbøger.
- JOW= Burke, John 1977: *Jowitt's Dictionary of English Law, vols. 1-2*. London: Sweet & Maxwell.
- LDBE= Adam, J.H. 1989: *Longman Dictionary of Business English*. Harlow: Longman/York Press.
- LMED= Watson, Owen (ed.) 1976: *Longman Modern English Dictionary*. Harlow/London: Longman.

### 7.2 Sekundære kilder

- Al-Kasimi, Ali M. 1977: *Linguistics and Bilingual Dictionaries*. Leiden: E.J. Brill.
- Bergenholtz, Henning 1990: Leksikographische Instruktionen für ein zweisprachiges Wörterbuch. I: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationswissenschaft* 43: 19-37.
- Bergenholtz, Henning & Dam, Helle & Henriksen, Torben 1990: Udarbejdelsen af en spansk-dansk juridisk ordbog: overvejelser og løsninger. I: *Hermes* 5: 127-136.
- Bergenholtz, Henning & Nielsen, Sandro 1988: Den hidtil største dansk-engelsk/engelsk-danske fagordbog. I: *Hermes* 1: 251-255.
- Bergenholtz, Henning & Poulsen, Sven-Olaf 1989: Leksikografi på HHÅ. Udvikling og perspektiver. I: *Handelshøjskolen i Århus 50 år. Festskrift i anledning af Handelshøjskolens 50-års jubilæum 31. august 1989*. Århus: Handelshøjskolen i Århus: 110-117.
- Cowie, Anthony P. 1983: English Dictionaries for the Foreign Learner. I: R.R.K. Hartman (ed.): *Lexicography: Principles and Practice*. London [etc.]: Academic Press: 135-144.
- Hartmann, Reinhard Rudolf Karl (ed.) 1983: *Lexicography: Principles and Practice*. London [etc.]: Academic Press.
- Hausmann, Franz Josef 1977: *Einführung in die Benutzung der neufranzösischen Wörterbücher*. Tübingen: Niemeyer.
- Hausmann, Franz Josef 1985: Lexikographie. I: Cristoph Schwarze & Dieter Wunderlich (eds.): *Handbuch der Lexikologie*. Königstein. Athenäum: 367-411.
- Hausmann, Franz Josef 1989: Wörterbuchtypologie. I: F.J. Hausmann & O. Reichmann & H.E. Wiegand & L. Zgusta (eds): *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol. 1*. Berlin/New York: de Gruyter: 968-981.

- Hausmann, Franz Josef & Reichmann, Oskar & Wiegand, Herbert Ernst & Zgusta, Ladislav (eds.) 1989: *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol. 1.* Berlin/New York: de Gruyter.
- Hausmann, Franz Josef & Reichmann, Oskar & Wiegand, Herbert Ernst & Zgusta, Ladislav (eds.) 1990: *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol. 2.* Berlin/New York: de Gruyter.
- Hausmann, Franz Josef & Wiegand, Herbert Ernst 1989: Component Parts and Structures of General Monolingual Dictionaries: A Survey. I: F.J. Hausmann & O. Reichmann & H.E. Wiegand & L. Zgusta (eds.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol. 1.* Berlin/New York: de Gruyter.
- Hupka, Werner 1989: Das enzyklopädische Wörterbuch. I: F.J. Hausmann & O. Reichmann & H.E. Wiegand & L. Zgusta (eds.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol. 1.* Berlin/New York: de Gruyter: 988-999.
- Kromann, Hans-Peder & Riiber, Theis & Rosbach, Poul 1984: Überlegungen zu Grundfragen der zweisprachigen Lexikographie. I: Herbert Ernst Wiegand (ed.): *Studien zur neuhochdeutschen Lexikographie V.* Hildesheim/New York: Olms: 159-238.
- Landau, Sidney I. 1984: *Dictionaries. The Art and Craft of Lexicography.* New York: Scribner Press.
- Malkiel, Yakov 1975: A Typological Classification of Dictionaries on the Basis of Distinctive Features. I: Fred W. Householder & Sol Saporta (eds.): *Problems in Lexicography.* Bloomington: Indiana University: 3-24.
- Moulin, André 1983: LSP Dictionaries for EFL Learners. I: R.R.K. Hartmann (ed.): *Lexicography: Principles and Practice.* London [etc.]: Academic Press: 144-152.
- Mugdan, Joachim 1989: Grammar in Dictionaries for Languages for Special Purposes (LSP). I: *Hermes* 3: 125-142.
- Nielsen, Sandro 1989: Kritisk oversigt over engelske og danske juridiske ordbøger. I: *Hermes* 2: 55-75.
- Nielsen, Sandro 1990: Contrastive Description of Dictionaries Covering LSP Communication. I: *Fachsprache/International Journal of LSP* 3-4: 129-136.
- Scholfield, P.J. 1980: Evaluating Selection Policy and Grammatical and Semantic Information in an EST Dictionary. I: *Fachsprache/International Journal of LSP, vol. 2:* 98-108.
- Wiegand, Herbert Ernst 1983: Was ist eigentlich ein Lemma? Ein Beitrag zur Theorie der lexikographischen Sprachbeschreibung. I: Herbert Ernst Wiegand (ed.): *Studien zur neuhochdeutschen Lexikographie III.* Hildesheim/New York: Olms: 401-474.
- Wiegand, Herbert Ernst 1988: Was eigentlich ist Fachlexikographie? Mit Hinweisen zum Verhältnis von sprachlichem und enzyklopädischem Wissen. I: Horst Haider Munske & Peter von Polenz & Oskar Reichmann & Reiner Hildebrandt (eds.): *Deutscher Wortschatz. Lexikologische Studien. Ludwig Erich Schmitt zum 80. Geburtstag von seinen Marburger Schülern.* Berlin/New York: de Gruyter: 729-790.
- Wiegand, Herbert Ernst 1989: Der Begriff der Mikrostruktur. Geschichte, Probleme, Perspektiven. I: F.J. Hausmann & O. Reichmann & H.E. Wiegand & L. Zgusta (eds.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol. 1.* Berlin/New York: de Gruyter: 409-462.
- Zgusta, Ladislav 1971: *Manual of Lexicography.* The Hague/Paris: Mouton.
- Zgusta, Ladislav 1988: *Lexicography Today. An annotated bibliography on the theory of lexicography.* Tübingen: Niemeyer.

## Kollokationsundersøgelse inden for engelsk merkantilt og økonomisk sprog med henblik på udarbejdelse af ordbogskoncept

Som grundlag for en ordbogsmodel opstilles i denne artikel en typologi, der tilstræber at afgrænse kollokationer fra henholdsvis idiommer og frie kombinationer, hovedsageligt ved hjælp af kvalitative kriterier. I denne forbindelse beskrives kollokationers semantiske kendetegn og de problemer, der er forbundet med at identificere kollokationer ved hjælp af frekvens. Endelig gives et forslag til den praktiske udformning af en kollokationsordbog.

### 1. Indledning

Kollokationer hører til de ordkombinationer, der typisk volder problemer i forbindelse med indlæringen af et fremmedsprog, fordi de ligesom idiommer er idiosynkratiske. Det vil sige, at de ikke kan konstrueres ud fra et sprogs generelle grammatiske og semantiske regler, som det er tilfældet med frie kombinationer, men er underlagt yderligere sprogbrugsrestriktioner.

Det er glædeligt, at de almene én- og tosprogede ordbøger i stigende grad er opmærksomme på kollokationerne, men på grund af de mange andre oplysninger, bliver der kun begrænset plads til dem - ofte kun i eksempelmaterialet. Som ordbogsbruger på jagt efter en bestemt kollokation, vil man derfor ofte blive frustreret over at bruge meget tid på mange forskellige opslag - mange gange uden at få svar.

Jeg har i en prisopgave beskæftiget mig med kollokationsbegrebet ud fra et leksikografisk perspektiv (Poulsen 1991 a). På grundlag af mine teoretiske overvejelser samt en analyse af 1000 kombinationer med den engelske term "inflation" vil jeg i det følgende tegne en profil af en kollokationsordbog, som jeg kunne ønske mig den.

### 2. Profil af en kollokationsordbog

Formålet med ordbogen vil være *sprogproduktion*; den skal hjælpe brugeren med at udtrykke sig på fremmedsproget ved hjælp af de idiosynkratiske forbindelser, der udgør den etablerede sprogbrug, og samtidig hindre interferens fra modersmålet<sup>1</sup>.

Ordbogen skal give et dækkende billede af opslagsordenes kollokationsmuligheder. - Man skal for det første kunne finde de *leksikalske funktioner*<sup>2</sup>, der er relevante<sup>3</sup> for det enkelte ord. Det vil sige, at de udtrykker en funktion, der er direkte relateret til ordets egen betydning eller de sammenhænge, det normalt forekommer i. Eksempler på sådanne funktioner er *høj grad* som i "fierce competition", "wide range" og "rapid growth", *at skabe/aktivere* som i "do business", "make a loss" og "compile a dictionary" og *standse/-*

*ophæve* som i "lift an embargo", "correct an imbalance" og "cure inflation". - Frie kombinationer, som ikke har en sådan direkte funktionsbestemt relation til opslagsordet, skal derimod ikke med, det gælder f.eks. "disappointing growth", "buy a dictionary" og "regret a loss".

For det andet skal brugeren præsenteres for de valgmuligheder, der er inden for de enkelte leksikalske funktioner. Hvis sprogbrugen altså tillader, at samme funktion for et givet ord, *basisordet*, udtrykkes ved at forbinde dette med en række andre ord eller *kollokater*, skal de alle anføres, sådan at man - f.eks. ud fra stilistiske kriterier - kan vælge netop den forbindelse eller *kollokation*, der bedst udtrykker det, man ønsker at sige.

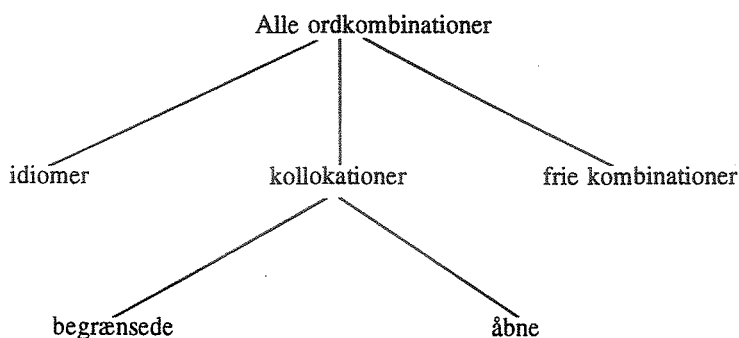
I nogle tilfælde er to ords indbyrdes kollokationsmuligheder begrænset til bestemte grammatiske former; f.eks. kan man tale om "runaway inflation" men ikke sige "inflation was running away". Sådanne *leksikogrammatiske restriktioner* skal også fremgå af ordbogen.

*Sammensatte og afledte former* er ikke selv kollokationer, som jo er syntagmer, men det vil være ønskeligt at medtage dem i en kollokationsordbog, da de ofte har de samme leksikalske funktioner og er underlagt leksikalske og leksikogrammatiske restriktioner. - Det er f.eks. nødvendigt at oplyse, at inflationsstigning på engelsk ikke hedder "\*inflation rise", men "rise in inflation", at sympatiaktion ikke hedder "\*sympathy action", men "sympathetic action", og at en forhandlingsløsning hedder "a negotiated settlement" og ikke "\*a negotiation settlement". Det er også en nyttig oplysning, at funktionen *høj grad* i forbindelse med "inflation" kan udtrykkes ved hjælp af præfikset "hyper-", ligesom *opposition* kan udtrykkes med "anti-", som i "anti-inflation policy".

Det er klart, at almene ordbøger ikke kan have plads til så fyldige oplysninger, så løsningen må være en specialordbog. Hvordan den kunne indrettes, vil jeg komme tilbage til under punkt 6.

### 3. Typologi og afgrænsning

De kollokater, der kan erstatte hinanden i en bestemt leksikalsk funktion, udgør et paradigme eller leksikalsk sæt. Et sådant sæt kan omfatte kollokater, hvis betydning er bundet til et enkelt eller et begrænset antal basisord (i tilfælde af begrænsede kollokationer), såvel som kollokater, der bruges i deres normale, selvstændige betydning (i tilfælde af åbne kollokationer). Det gælder altså om at afgrænse begge disse kollokationstyper samlet fra idiomers på den ene side og frie forbindelser på den anden<sup>4</sup>:



(Poulsen 1991 a: 20)

Overgangen mellem de forskellige typer af ordkombinationer er glidende, og kollokationer kan beskrives som en del af en skala, der går fra minimalt til maksimalt variable forbindelser. Det kan illustreres ved hjælp af følgende forbindelser med "bill" ("regning"):

IDIOMER		fill }		
		fit }		
		foot		
	KOLLOKATIONER {	BEGRÆNSEDE	pick up }	the bill
	}	ÅBNE	meet	
			pay	
			discuss	
FRIE KOMBINATIONER			tear up }	
			receive }	
			etc	

I den mindst variable ende af skalaen, har vi "fill" eller "fit the bill" som eksempler på *idiomer*. De tillader kun meget lidt variation, og i forbindelse med sprogproduktion fungerer de som en helhed ("unit of speech production"), idet de enkelte ord har mistet deres selvstændige betydning.

I midten kommer *kollokationerne*, der kan betragtes som en slags halvfabrikata i sproget, idet det er et kendetegn, at kollokatet kan skiftes ud, uden at basisordet ændrer betydning (Cowie 1981: 224).

Overgangen mellem de begrænsede kollokationer ("foot/pick up the bill") og de mere åbne kollokationer ("settle/meet/pay the bill") er glidende, og afhænger af, om kollokatet forekommer med uændret betydning i få eller mange forbindelser.

Fra et sprogproduktionssynspunkt er dette ikke så vigtigt; udgangspunktet er basisordet, og det, der betyder noget, er at få præsenteret hele spektret af udtryksmuligheder for de funktioner, der er relevante for det pågældende ord. - Hvis f.eks. de mest åbne kollokationer ikke er med, kan det skabe det fejlagtige indtryk, at de er mindre acceptable. Som regel vil de mest åbne kollokationer også være de mest neutrale fra et stilistisk synspunkt; en udeladelse vil derfor medføre, at dette formalitetsniveau mangler.

*De frie kombinationer* ("discuss, tear up, receive the bill", etc.) adskiller sig fra kollokationerne ved ikke at være underlagt sprogbrugsrestriktioner. De kan konstrueres ud fra sprogets generelle syntaktiske og semantiske regler, og der er ikke tale om et mønster, hvor et basisord vælger sig en bestemt partner ud fra paradigmer eller leksikalske sæt, der er fastlagt af sprogbrugen.

#### 4. Semantiske kendetegn

Kollokatet i begrænsede kollokationer udviser nogle karakteriske indholdstræk, som det er påpeget hos Aisenstadt (1979: 71) og Cowie (1989).

I én type kollokationer, er der tale om "*tomme*" *verber* (delexical verbs); indholdet af den handling, der er tale om, defineres udelukkende af objektet, og disse kollokater har

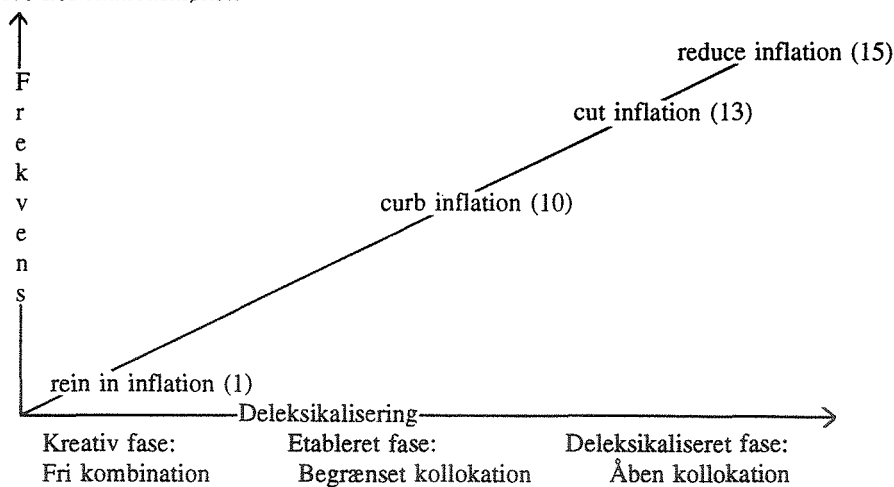
nærmest kun en grammatisk verbaliseringsfunktion, f.eks. "make a deal", "do business", "take action".

I en anden type, er der derimod tale om meget indholdsmættede kollokater, som er bundet til bestemte basisord på grund af deres meget specialiserede, nærmest *tekniske indhold*. Eksempler herpå er: "Engage a gear", "pass a law", "levy a tax", og "service a debt".

Ofte har kollokater et *metaforisk indhold*, som i "lift/impose a restriction", "pursue a policy", "go into business", "rapid growth" og "tight control".

Det er imidlertid først, når de udgør en del af den etablerede sprogbrug, at metaforer kan defineres som kollokationer. I den første, kreative fase, må de anses for at være frie kombinationer, selvom de falder inden for et af de funktionsbestemte leksikalske sæt. Jo mere etableret en metafor bliver, desto mere falmer billedet, således at de hyppigst anvendte billeder til sidst ikke længere opfattes metaforisk (Sinclair 1987: 323). Denne gradvise nedslidning af det metaforiske indhold ("deleksikalisering") illustreres nedenfor med nogle eksempler, hentet fra en samling omfattende 1000 kombinationer med "inflation".

Talmaterialet er naturligvis ikke omfattende nok til at bevise noget, men peger dog på en sandsynlig sammenhæng mellem frekvens og deleksikalisering. De alternative kombinationer, der kan forekomme inden for samme leksikalske funktion, befinder sig på forskellige udviklingstrin, og det ser ud til, at deres metaforiske "friskhed" og den hyppighed, med hvilken de forekommer, er omvendt proportionale. Dette er i hvert fald tendensen i de 1000 kombinationer med "inflation", idet de mest åbne kollokationer også var de hyppigste. Dette bekræfter efter min mening, at de åbne kollokationer hører med til beskrivelsen af et givet ords kollokationsmønster.



(Poulsen 1991:73).

## 5. Betydningen af frekvens

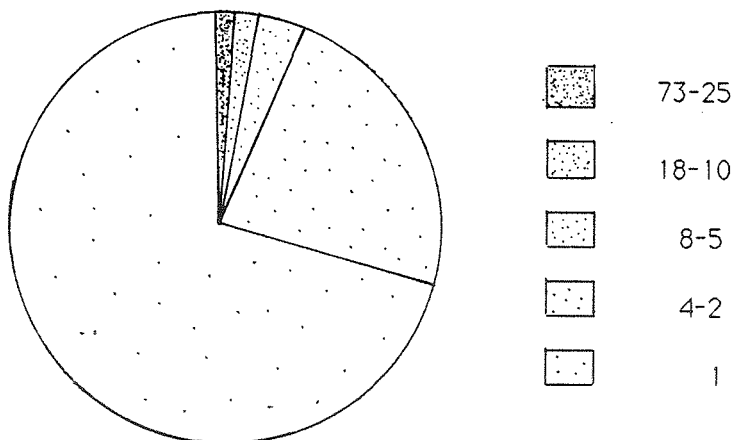
Da kollokationer jo er udtryk for sprogbrugsmønstre, er det klart, at frekvens eller regelmæssig forekomst spiller en rolle. Som det fremgår af det foregående, er det nogle gange nødvendigt at tage stilling til, om en metaforisk forbindelse, som "rein in inflation" er tilstrækkeligt etableret til at være en del af sproget, eller om den er udtryk for en kreativ

variation af den traditionelle hestemetafor for "inflation", som vi kender fra forbindelser med "runaway", "galloping" og "curb" (Poulsen 1991 a:73).

Et problem med at bruge frekvens som kriterium for, om en forbindelse er en kollokation eller en fri forbindelse, er imidlertid, at leksikalske mønstre til forskel fra grammatiske mønstre har en meget "lav generalitet" (Halliday 1966:159); de kan kun konstateres på grundlag af meget store tekstmængder<sup>5</sup>.

Den omtalte analyse af 1000 eksempler med "inflation" bekræfter dette. Der var hele 446 forskellige kombinationer, og nedenstående opstilling viser, hvor hyppigt de forekom:

Kombinationernes hyppighed	Antal kombinationer	Procentdel af alle kombinationer
73-25	5	1.12
18-10	9	2.02
8-5	19	4.26
4-2	102	22.87
<u>1</u>	<u>311</u>	<u>69.73</u>
73-1	446	100.00



311 eller næsten 70% af dem optrådte kun en enkelt gang, heriblandt metaforiske kombinationer som "rekindle inflation", "reignite inflation", "runaway inflation" samt "inflationary spiral", der ud fra min samlede sproglige erfaring er etablerede. Andre kombinationer, der synes typiske, forekom kun to til fire gange, herunder "galloping inflation", og "fight inflation". Denne gruppe udgjorde knap 23%.

Kun godt 3% forekom mere end 10 gange. De allerhyppigste kombinationer var "inflation rate" (73 gange) og "rate of inflation" (33 gange), hvor "rate" har den grammatiske funktion at gøre "inflation" tælleligt ("count noun"). - Disse to sammensætninger udgjorde alene over 10% af samtlige eksempler. Også afledninger som "anti-inflation" (30 gange) og



"hyper-inflation" (18) var blandt de hyppigste sammen med åbne kollokationer som "high(er) inflation" (41), "reduce inflation" (15) og kombinationer med en form af "fall"(17).

Nogle frie kombinationer forekom ret hyppigt, men de åbne kollokationer var de hyppigste, sammen med afledninger og sammensætninger. Det store flertal af kombinationer, der kun forekom én gang, omfattede åbne og begrænsede kollokationer såvel som frie kombinationer.

At bruge frekvens som eneste kriterium giver altså problemer. - Som påpeget af Hausmann (1985: 125), er der kollokationer, der ikke optræder særlig hyppigt, men som alligevel er "til rådighed":

"Viele Kollokationen sind nicht frequent aber dennoch verfügbar"

Den lave frekvens kan for eksempel skyldes et højt formalitetsniveau - således vil kollokationen "to perpetrate a crime" sandsynligvis ikke forekomme nær så ofte som "to commit a crime". På den anden side vil mange frie kombinationer kunne forekomme hyppigt, fordi de refererer til en meget almindelig situation, snarere end på grund af et særligt sprogbrugsmønster, f.eks. "to buy a dictionary, shoes, flowers, ... etc."

Jeg mener altså ikke, at frekvens kan erstatte kvalitative kriterier (tilhørsforhold til leksikalske sæt med relevante leksikalske funktioner) som en generel metode til afgrænsning af kollokationer. Hvis man har et egnet korpus til rådighed, er der derimod ikke noget i vejen for at inddrage frekvensaspektet i sin vurdering af konkrete spørgsmål - f.eks. om en metaforisk kombination som "rein in inflation" er en kollokation eller en fri kombination, eller om "survey" kollokerer med et eller flere af verberne "carry out, do, perform, make og conduct".

## 6. En ordbogsmodel

Det jeg ønsker er altså en specialordbog, der udelader de almene ordbøgers generelle oplysninger for i stedet at give fyldige oplysninger om ordenes kollokationsmønstre, og som primært er baseret på kvalitative udvælgelseskriterier.

*The BBI Combinatory Dictionary of English* (Morton Benson et al 1986) er en sådan ordbog. Den indeholder mange kollokationer, som man ikke før har kunnet slå op, og er derfor en velkommen fornyelse.

Det viser sig imidlertid vanskeligt at dække alle sprogområder i én ordbog, og især inden for fagsprogsområderne, er dækningen meget ujævn.

En anden mangel er det, at forfatterne ikke har en klar holdning til afgrænsningsproblemet mellem kollokationer og frie forbindelser, hvilket fører til inkonsekvens med hensyn til behandlingen af åbne kollokationer (Poulsen 1991 b).

Det er nok urealistisk at forvente, at én specialordbog kan dække et helt sprog, og jeg mener derfor, at man skal koncentrere sig om et område ad gangen. Da det jo er de fagsproglige områder, der lider mest under pladsmangelen i de almene ordbøger, kunne man passende begynde her - for eksempel med merkantilt og økonomisk sprog.

Efter min mening vil en énsproget ordbog på fremmedsproget være bedst egnet til formålet, idet den kan præsentere det pågældende sprogs kollokationer uden hensyn til sprogbrugsmønstre på modersmålet. Samtidig vil den være velegnet til at aktivere en eventuel passiv viden.

Målgruppen for en sådan ordbog vil være "advanced learners", der normalt har et omfattende passivt kendskab til ordenes kombinationsmuligheder, og hvis viden om de generelle syntaktiske og semantiske regler er tilstrækkelig til, at de kan have glæde af en specialordbog.

En *alfabetisk makrostruktur* vil gøre oplysningerne lettilgængelige, men udvalget af opslagsord bør være baseret på en systematisk gennemgang af det pågældende fagområde, og det bør oplyses, hvilke delområder, der er dækket. - Kollokationerne vil kun være at finde under basisordet, som hyppigst vil være et substantiv, idet verber og adjektiver dog kan være basisord i forbindelse med adverbier og præpositioner.

*Mikrostrukturen opdeles systematisk* efter de leksikalske funktioner, der er relevante for det enkelte opslagsord, og også syntaktiske mønstre i form af leksikogrammatiske restriktioner bør fremgå her. De funktioner, der medtages, vil variere for de forskellige opslagsord, og de bør have umiddelbart forståelige, engelske betegnelser, idet det ikke forventes, at den normale bruger har tålmodighed til at sætte sig ind i et indviklet kodesystem.

En ordbog, som den der er skitseret her, tror jeg kan være en god hjælp til at undgå de mange fælder, der opstår, når man skal finde rundt i et fremmedsprogs indviklede sprogbrugslaboratorium, samtidig med at man må frigøre sig fra de indgroede vaner, man bringer med sig fra sit modersmål.

## Noter

1. Jeg har kun medtaget kombinationer af leksikalske enheder. Valensbetingede kombinationer af leksikalske enheder og grammatiske strukturer som f.eks. verbalmønstre udgør selvfølgelig også en del af sprogbrugen, men er efter min erfaring dækket ind i de almene ensprogede engelske ordbøger.

2. Begrebet "leksikalsk funktion" stammer fra Sovjetunionen, hvor det blev anvendt af Apresyan, Melčuk og Žholkovsky i forbindelse med deres arbejde på en ny type ordbog: *The Explanatory and Combinatory Dictionary of Modern Russian* (Apresyan et al 1969: 1-19). Den engelske definition lyder som følger: "A standard lexical function is a meaning relation between a key word (or word combination) Co and other words and word combinations Ci, which meets the following three requirements: (1) this relation occurs in a sufficiently great number of word pairs, i.e. it manifests itself through many different Co; (2) this relation has diverse means of expression, i.e. the number of different Ci in the language is rather great; (3) the choice of the right Ci for the expression of a given relation with a given Co is determined, as a rule, by the Co itself (sometimes under supplementary conditions)."

3. Viehweger (1982:27) peger på tre typer kombinationer med "Blume", der udviser sprogbrugsrestriktioner, nemlig med verber der udtrykker en plantes typiske funktioner eller egenskaber, med verber der udtrykker menneskelige handlinger i forbindelse med blomster, og endelig med verber der udtrykker sådanne handlinger i typiske sociale sammenhænge.

4. Andre typologier skelner ikke mellem åbne kollokationer og frie kombinationer, men slår disse begreber sammen under én af betegnelserne. F.eks. opdeler Aisenstadt (1979:71) alle ikke-idiomatiske ordkombinationer i "restricted collocations" og "free phrases", og Cowie definerer i introduktionen til bind 2 af "The Oxford Dictionary of Current Idiomatic English" betegnelsen "open collocation" så den omfatter både åbne kollokationer og frie kombinationer (Cowie et al. 1983).

5. Halliday (1966:159) mente, at det ville være nødvendigt med en tekst på 20 millioner løbende ord eller mellem 1.500 og 2.000 timers samtale (1966:159) på grund af de leksikalske mønstres "lave generalitet": "... since lexical patterns are of low generality, they appear only as properties of very large samples".

## Litteratur

- Aisenstadt, E. 1979. Collocability restrictions in dictionaries. I: *ITL Review of Applied Linguistics* 43-46: 71-74.
- Apresyan, Yu. D., I.A. Melčuk and A.K. Žolkovsky. 1969. Semantics and lexicography: towards a new type of unilingual dictionary. I: F. Kiefer, red. *Studies in Syntax and Semantics*:1-33. Dordrecht - Holland.
- Benson, Morton, Evelyn Benson og Robert Ilson. 1986. *The BBI Combinatory Dictionary of English: A Guide to Word Combinations*. Amsterdam/Philadelphia.
- Cowie, A.P. 1981. The treatment of collocations and idioms in learners' dictionaries. I: *Applied Linguistics* 23.
- Cowie, A.P., R. Mackin og I.R. McCaig. 1983. *Oxford dictionary of current idiomatic English, Indledning til bind 2, 0.1*. Oxford.
- Cowie, A.P. 1989. The language of examples in English learners' dictionaries. I: G. James, red. *Lexicographers and their works*. Exeter.
- Halliday, M.A.K. 1966. Lexis as a linguistic level. I: *In Memory of J.R. Firth*.
- Hausmann, Franz Josef. 1985. Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. I: *Akten des Essener Kolloquiums zur Grammatik im Wörterbuch, 28.-30.6. 1984*. Tübingen.
- Poulsen, Sonja. 1991 a. *Word Combinations as Lexical Building Blocks. A theoretical framework and a possible concept for a specialized dictionary of English collocations for text production*. Besvarelse af prisopgave ved Odense Universitet. Ikke publiceret.
- Poulsen, Sonja. 1991 b. Anmeldelse af *The BBI Combinatory Dictionary of English: A Guide to Word Combinations*. I: *Hermes, Tidsskrift for sprogforskning, nr. 7 - 1991*. Århus.
- Sinclair, J. McH. 1987. Collocation: a progress report. I: R. Steele, T. Threadgold, udg.: *Language topics, essays in honour of Michael Halliday (1-2)*, 319-331. Amsterdam, Philadelphia.
- Viehweger, Dieter. 1982. Die Darstellung semantischer Vereinbarkeitsbereiche zwischen lexikalischen Elementen im einsprachigen Wörterbuch des Deutschen. I: E. Agricola, J. Schildt, D. Viehweger, udg.: *Wortschatzforschung heute. Aktuelle Probleme der Lexikologie und Lexikographie*. Leipzig.

## Om opstilling af synonymmer i synonym- og betydningsordbøger

Artiklen handler om præsentationen af synonymmer i synonym- og betydningsordbøger. Der gennemgås forskellige principper for listeformet opstilling af synonymmer med hensyn til følgende forhold: Gruppering af synonymmer med fast eller varierende gruppestørrelse, pakning ved hjælp af samleartikler, henvisningsfunktioner og markering af relationer inden for synonymgruppen. Markering af relationelle forskelle ses i sammenhæng med integration af synonymmer i betydningsordbøger.

### Synonymmer i synonym- og betydningsordbøger<sup>1</sup>

De to yderpunkter i præsentationen af synonymmer, den sammenlignende fremstilling og den listeformede opregning, findes begge i synonymordbøger og i betydningsordbøger, der har synonymmer som eksplicit anført kategori.

Ved den første præsentationsmåde sammenlignes synonymmer med hensyn til ligheder og forskelle i betydning og anvendelse. Eksempler på ordbøger med vægt på sådan diskriminerende synonymik er *Webster's Dictionary of Synonyms* og *The Penguin Modern Guide to Synonyms and Related Words*. I Danmark er typen repræsenteret med *Dansk Synonymik* (sidste udgave 1872). Blandt de betydningsordbøger, som anfører synonymmer som eksplicit angivet kategori, findes præsentationsformen i bl.a. *Longman Dictionary of the English Language*. Synonymbeskrivelsen under opslaget *subject* indledes her sådan:

- (1) *Subject, object, topic and theme* can all mean "that to which one's attention is directed". *Subject* and *object* are sometimes interchangeable <a *subject*/an *object* of study>, but *subject* emphasizes [...]

Det er pladskrævende at redegøre for synonymi på denne måde, og udvalget af synonymmer er da også yderst selektivt i de betydningsordbøger, hvor denne præsentationsform forekommer.

Synonymordbøger med den anden og mere udbredte præsentationsmåde, hvor synonymmer opregnes i listeform, er fx *Webster's Collegiate Thesaurus*, *The Oxford Thesaurus*, *Duden Sinn- und sachverwandte Wörter*, *Ord för ord*, *Dansk synonymordbog* og *Politikens Synonymordbog*.

Når der i bøger som disse benyttes et bredt synonymibegreb ved grupperingen af ord, bygger præsentationsformen i høj grad på brugerens kompetens (og typen fungerer da også kun med henblik på produktion og ikke med henblik på forståelse af tekst). Selv om listede synonymmer kan være kommenteret med markører for stil, brugsområde etc., og selv om der eventuelt skelnes mellem forskellige grader af synonymi, således at fjernere synonymmer er markerede, forudsættes det, at den tilstrækkelige viden om ordenes betydning og brug kan

aktualiseres hos brugeren selv, eller at denne (hvis han ellers erkender sit behov derfor) supplerer sin viden i en betydningsordbog. Jf. *Politikens Synonymordbog* s. 8 om denne type: "det overlades til læseren at skønne over indbyrdes ligheder og forskelle i ordenes anvendelse"; jf. også opfordringen til læseren i denne og andre ordbøger om at søge støtte i en betydningsordbog.

Eksempler på betydningsordbøger med synonymer præsenteret i listeform er *Duden Bedeutungswörterbuch* og *Politikens Nudansk Ordbog*, som har anført længere rækker af synonymer ved de enkelte ordbetydninger, og *Collins Concise Dictionary and Thesaurus*, som er en kompilering af en betydnings- og synonymordbog.

I *Collins Cobuild English Language Dictionary* og *Svensk ordbok* er der ved de enkelte betydninger anført ganske få betydningsbeslægtede ord (strikte synonymer, antonymer og henholdsvis hyperonymer og fjernere betydningsbeslægtede ord); udvalget fungerer her ikke som i en gængs synonymordbog (jf *Svensk ordbok* s. xvii), men har snarere karakter af et makrostrukturelt supplement til definitionen. Tilsvarende knytter *Ordbog over det Danske Sprog* synonymer til definitionerne ("til selve forklaringen [...] føjes de vigtigste synonymer", bind 1 s. xxxviii), markeret som selvstændigt felt, idet de altid følger efter en længere parafrase, adskilt fra denne ved semikolon.

I det følgende gælder fremstillingen synonymer opført i listeform i ordbøger med flere synonymer.

### Synonymer præsenteret i listeform

Et forhold, som adskiller præsentationen af listede synonymer i forskellige ordbøger, er, om den enkelte synonymgruppe fremtræder som konstant mængde eller ej i forhold til alle medlemmer i gruppen.

En gruppe af betydningsbeslægtede ord *a, b, c, d, e, f* kan fungere som fast, genkommende gruppe for hvert af medlemmerne, når disse står som opslagsord:

- (2) A: b, c, d, e, f  
E: a, b, c, d, f

Dette princip kan udmøntes som i *Webster's Collegiate Thesaurus*, hvor gruppen er præsenteret ved hvert opslagsord, fx:<sup>2</sup>

- (3) BARREL: CASK, butt, hogshead, keg, pipe, tun  
CASK: barrel, butt, hogshead, keg, pipe, tun

Eller det kan realiseres sådan, at der benyttes samleartikler, hvortil der henvises, således i *Duden Sinn- und sachverwandte Wörter*, i *Politikens Synonymordbog* (hvorfra eksempel 4) samt i *Politikens Nudansk Ordbog* (\* er henvisningsmarkør):

- (4) ADSKILLE: sondre\*  
DIFFERENTIERE: sondre\*  
SONDRE: adskille, holde adskilt, differentiere, distingvere, gøre forskel (på), gøre (sætte) skel (imellem), skelne, uddifferentiere, holde ude fra hinanden

Ved begge opstillingsformer kan ordene i gruppen være ligeværdige som opslagsord for gruppen. Eller opstillingen kan i en vis grad afspejle relationelle forskelle inden for gruppen, idet et af ordene har status som hovedopslagsord og er det, der som "Leitwort" eller "Grundsynonym" præger gruppens sammensætning og som sådant er det mindst markerede eller det hyppigste.<sup>3</sup>

I *Webster's Collegiate Thesaurus* skelnes der således inden for den enkelte synonymgruppe mellem ord, som fungerer som henholdsvis "main entry" og "secondary entry". I den viste synonymgruppe (eksempel 3) er *cash*, som er det ord, der har den mest generelle betydning, hovedopslag, og ved dette er angivet den betydningskerne, som er fælles for gruppen, ligesom forskellig anden information er anført her (bl.a. præsenteres rækker på mere end ti synonymer kun i deres helhed under hovedopslagsordet).

Tilsvarende er det i ordbøger med samleartikler placeringen af disse, der kan være udtryk for relationelle forhold i gruppen. Således placeringen af grupper med hyponymrelationer under et hyperonym, fx, fra *Politikens Synonymordbog*, *LÅG: dække, dæksel, hætte, kapsel, klap, skruelåg*.

En gruppe betydningsbeslægtede ord kan også have en mere eller mindre singular forekomst, idet den ikke fremtræder konstant i forhold til gruppens medlemmer, når disse står som opslagsord, men ændrer udseende under forskellige opslagsord:

- (5) A: b, c, d, e  
E: a, b, f

(idet det her forudsættes, at der under *A* og *E* ikke også optræder grupper identiske med henholdsvis *E*- og *A*-grupperne).

Således har *Dansk synonymordbog* disse synonymgrupper, hvori *adskille, skelne* og *sondre* indgår med fællesbetydningen 'skelne':

- (6) ADSKILLE: sætte skel imellem, sondre, skelne, kende fra hverandre  
SKELNE: sondre, adskille  
SKILLE: skelne, sondre\*  
SONDRE: distingvere, holde ude fra hinanden, holde adskilt, opløse, skelne\*, skille

Her er fx *skelne* relateret til *SKILLE*, mens der ikke er nogen relatering mellem *SKELNE* og *skille*, og *skelne* er relateret til *distingvere* og *holde ude fra hinanden* under *SONDRE*, men ikke under *ADSKILLE* og *SKELNE*.

Samme forhold præger flere ordbøger, således også *Ord för ord* og *Collins Concise Dictionary and Thesaurus* (fx *PESSIMISM: mörk livssyn l. livsåskådning, mörksyn, svartsyn, (livs)bitterhet, världsförakt, jfr melankoli\*, negativism, skepticism* og *SVARTSYN: pessimism, melankoli, negativism*).

Ved varierende gruppestørrelse er det det enkelte opslagsord, der bestemmer gruppens sammensætning, og relationerne i gruppen må læses primært i retningen fra opslagsordet til den øvrige del af gruppen. Dette er udtrykt eksplicit i *The Oxford Thesaurus* s. x: "it is the function of this book to provide synonyms for the main entries, not for the other synonyms that may be listed".

Som redaktionel fremgangsmåde kan det, at der tages udgangspunkt i det enkelte opslagsord, medføre, at anførelse og afgrænsning af synonymer under forskellige opslagsord

bliver inkonsistent: Sæt med ensartede indbyrdes relationer kan blive flossede. Og en symmetrisk relation mellem *a* og *e* kan optræde under *A*, men mangle under *E*:

- (7) A: *b, c, d, e, f*  
E: *b, f*

Men variation i gruppestørrelse kan også være udtryk for, at opslagsordene styrer singulære sæt på en mere principiel måde. Dette kan være tilfældet, hvis grupperingen konstitueres af et synonymibegreb, som tillader ensidig implikation i form af hyperonyme, men ikke hyponyme relationer (her vil fx *skruelåg* have tilordnet *låg*, men den omvendte tilordning vil ikke forkomme). Læst på denne måde vil

- (7') A: *b, c, d, e, f*  
E: *b, f*

være udtryk for relationelle forskelle, hvor *b, e, f* er hyperonymer i gruppen *a, b, c, d, e, f*, og *a, c, d* vil da ikke figurere under de hyperonyme opslagsord *B, E, F*. Relationerne i (7') kan her kun læses envejs fra opslagsordet, man kan ikke læse inde fra rækken, dvs. fra *b, e* og *f* til opslagsordet eller til de andre ord i rækken.

Et sådant princip synes til dels at ligge bag den asymmetriske præsentation i *The Oxford Thesaurus*. Fx anføres her under *BITTER 2* bl.a. *distasteful, unsavoury, obnoxious, disagreeable, nasty, unwelcome*, af hvilke ingen som opslagsord har anført *bitter* som synonym.

Læst som relationel markering vha. gruppering kan (7') konverteres til en direkte markering af de relationelle forskelle, således at gruppestørrelsen er intakt som i (2), mens det så er markeringerne, der varierer, efter hvilket ord der er opslagsord ( $\uparrow$  og  $\downarrow$  står for henholdsvis hyperonymi og hyponymi):

- (8) A:  $\uparrow b, c, d, \uparrow e, \uparrow f$   
E:  $\downarrow a, b, \downarrow c, \downarrow d, f$

(eller sammenpakket med henvisning til samleartikler: A: *c, d,  $\uparrow e^*$*  og E:  $\downarrow a^*, b, f$ ). Her kan man igen læse begge veje, altså også inde fra rækken af ord under opslagsordet og ud. Relationel markering omtales videre nedenfor, eksempel (27).

### Pakning: samleartikler, henvisningsformer

Når der er tale om større grupper af synonymer (hvis synonymibegrebet er bredt), kan pladshensyn tale for at pakke synonymerne i samleartikler, hvortil der henvises fra andre ord. Især hvis der også angives forbindelser til fjernere betydningsbeslægtede grupper, vil det være nødvendigt at henvise i stedet for at præsentere grupperne i deres helhed ved hvert opslagsord. Fx ville en gennemført pakket ordbog som *Politikens Synonymordbog* vokse til næsten det tredobbelte omfang, hvis artiklerne fremtrådte i upakket form.

Pakning forudsætter etablering af ordgrupper, som fungerer konstant med samme afgrænsning i forhold til hvert af medlemmerne i gruppen, når disse står som opslagsord.

Pakning af grupperne under *adskille* og *sondre* i eksempel (6) ovenfor vil således ikke være mulig, med mindre grupperne gøres ens.

Det er ved pakning også selvfølgelig (om end det ikke altid forekommer), at det opslagsord, hvorfra der henvises, figurerer i samlegruppen under det opslagsord, hvortil der henvises. I sidste ende ville der jo ikke være nogen samlegruppe, hvis henvisningsordene ikke stod i den. Og ved én form for artikelstrukturering med flere grupper er det påkrævet, for at henvisning kan fungere.

Henvisning fra *E* til *A*:

- (9) A: b, c, d, e, f  
E: a\*

læses som handleanvisning sådan: 1) gå til *A*, 2) ord relateret til *E* er *A*, *b*, *c*, *d*, [*e*], *f*.

En sådan henvisningsfunktion med to trin er imidlertid ikke tilstrækkelig i de tilfælde, hvor der under et opslagsord er flere grupper, som ikke er adskilt ved et mærke i form af et tal eller lignende, men kun ved semikolon eller et andet lignende tegn. Fx:

- (10) A: b, c; d, e, f  
E: a\*

Her er det nødvendigt at matche opslagsordet for henvisningen med et ord i en af grupperne, altså: 1) gå til *A*, 2) match *E* med *e*, 3) ord relateret til *E* er *A*, *d*, [*e*], *f*, dvs. opslagsord plus den semikolonafgrænsede gruppe, hvori det matchede ord *e* står.<sup>4</sup> Eksempler herpå fra *Duden Sinn- und sachverwandte Wörter* er:

- (11) LÖFFEL: Esslöffel, Suppenlöffel; Kochlöffel, Ruhrlöffel; Gemüselöffel, Sosselöffel, Saucelöffel; Kaffeelöffel, Teelöffel, Mokkalöffel, Dessertlöffel, Eierlöffel  
TEELÖFFEL: Löffel\*

Og fra *Politikens Nudansk Ordbog*:

- (12) SKELNE: sondre, holde ude fra, adskille; skimte, øjne, ane  
SONDRE: skelne\*

Et eller flere ord kan høre til i forskellige grupper under et opslagsord:

- (13) A: b, c, d; b, c, e, f

Man kan naturligvis bortspare sådanne gentagelser (jf. *Dansk synonymordbog* s. 7 "visse ord burde egentlig stå i flere grupper, men af pladshensyn er sådanne gentagelser sparet", således også *Ord för ord* s. xiii). En anden måde at spare på er at pakke vha. markering af, hvad der er fælles for flere grupper, fx ved brug af kolon:

- (14) A: b: c; d, e, f



Henvisningen *E: a\**, gælder så her gruppen *A, b, d, [e], f*. Denne form for pakning inden for en artikel er brugt i *Politikens Nudansk Ordbog*, fx:

(15) TILBAGE: igen: retur; til rest, tilovers; - bagud

som er en pakning af grupperne *tilbage, igen, retur* og *tilbage, igen, til rest, tilovers* og, adskilt ved tankestregen som en gruppe, der kun har opslagsordet fælles med de andre, *tilbage, bagud*.

Henvisning fra et opslagsord til en samlegruppe som i (9) er naturligvis ensrettet. Men der kan også være tale om, at grupper af synonymmer er relateret til hinanden, og her vil henvisning være symmetrisk, med mindre henvisning bruges - ligesom variation i gruppestørrelse i (7) - til at signalere ensidige relationer.

Henvisning fra en gruppe til en anden markerer normalt et skifte eller en graduering i betydningsrelationer i forhold til dem, der konstituerer gruppen. Henvisningsordet kan da have en markør (jf. eller et andet adskilletegn, fx semikolon), der angiver relationsforskydningen:

(16) A: b, c, d, jf. e\*  
E: f, jf. a\*

Eller henvisningsordet kan stå som medlem af gruppen uden særlig markering, hvis det tilhører gruppen på samme relationelle niveau som de øvrige medlemmer:

(17) A: b, c, d, e\*  
E: a\*, f

I (17) kan der være tale om, at henvisningsordet er polysemnt og altså hører til de forskellige grupper i forskellige betydninger. (17) vil her kunne svare til en opstilling *A: e\** og *E: a, b, c, d; f* (på nær, at det ved denne opstilling ikke er angivet, at der er nogen relation imellem *a, b, c, d, e* og *a, e, f* ud over fællesordene *a* og *e*). En anden og tydeligere præsentation af (17) med denne læsning vil være:

(18) A: b, c, d, e, jf. e\*  
E: a, f, jf. a\*

Men (17) kan også blot være udtryk for, at henvisningsordet hører til i grupper baseret på en intern associering, der er så løs, at det kan optræde på niveau med medlemmerne i begge grupper.

I *Duden Sinn- und sachverwandte Wörter* relateres grupper med en henvisningsform som i (16). Her står henvisningsord, der forbinder grupper, adskilt fra den gruppe, der henvises fra, med semikolon som forskydningsmarkør. I *Politikens Synonymordbog* bruges begge henvisningsformer (16) og (17), den første markeret med *jf.* plus kursiv skrift, den sidste med kursiv skrift alene. Henvisning mellem grupper er gennemført symmetrisk i disse to ordbøger.

En særlig form for graduering er henvisning fra hyponymer i en bredt konstitueret gruppe til mere snævert konstituerede grupper, der enten kompletterer eller afgrænser disse. En sådan relateringsform forekommer i *Duden Bedeutungswörterbuch*, fx:

- (19) BERUFUNG: Rekurs; Einspruch\*  
EINSPRUCH: Berufung\*, Beschwerde, Demarche, Einsprache, Einwand, Einwendung,  
[...] Rekurs

Her er henvisningstyperne (16) og (17) brugt asymmetrisk på den måde, at (16) (*BERUFUNG: ...; Einspruch\**) henviser til en bredere konstitueret gruppe under et hyperonymt opslagsord, mens (17) (*EINSPRUCH: ..., Berufung\**) henviser til afgrænsning af hyponymer i en snævrere konstitueret gruppe.

*The Oxford Thesaurus* benytter alene henvisningsformen (17). Henvisningstegnet angiver her, at det pågældende ord også er opslagsord i ordbogen og ikke nødvendigvis, at der er tale om en relationel graduering (sådan som det antydes i introduktionen s. xiii, nemlig at henvisningstegnet "indicates that the word also appears as a main entry at which the reader may find further useful synonyms"). Fx giver henvisningen *THANKFUL: grateful\* ...* ingen nye synonymer, idet synonymerne til *GRATEFUL* er indeholdt i synonymrækken til *THANKFUL*.

Det kan tilføjes, at henvisning mellem grupper ikke nødvendigvis behøver at være udtryk for graduering af relationer. Man kan således tænke sig en gruppe *a, b, c, d, e, f* fordelt på flere samlegupper, fx *A: b, c, d, e\** og *E: a\*, f*, uden at der dermed angives nogen relationel forskel. En sådan præsentation vil dog være unødigt komplicerende.

Ved henvisning mellem grupper kan disse til gengæld indgå i åbne systemer, hvor man kan bevæge sig længere og længere væk fra udgangspunktet gennem forskellige betydningsfelter. Hvis man fx slår op på *bange* i *Politikens Synonymordbog*, fører én rute af henvisninger således gennem følgende opslagsord:

- (20) BANGE → FEJ → KARAKTERLØS → USTADIG → HVILELØS → VIRKELYSTEN  
→ HANDLEKRAFTIG → ARBEJDSDYGTIG

Uden brug af henvisninger, må sådanne netværk lukkes efter første trin. Dette gøres i *Webster's Collegiate Thesaurus*, der anfører grupper af "related words" i de enkelte artikler uden at benytte henvisning. (Man kan se dette som etablering af bredere faste grupper med markeret niveaudeling i gruppen, for så vidt at disse fjernere relaterede grupper ikke varierer i forhold til forskellige opslagsord).

### Kommentarer og relationsmarkører

Listeformede opstillinger af betydningsbeslægtede ord kan have kommentarer, som på forskellige måder specificerer forholdet mellem ordene.

*Duden Sinn- und sachverwandte Wörter, Ord för ord og Politikens Synonymordbog* har flere forskellige systemer af sådanne kommentarer, bl.a. med angivelse af betydningsområde, fx:

- (21) ELEKTROGRAMM: Elektrodiagnose; *des Herzens*: Elektrokardiogram, [...]; *der Herzrandbewegungen*: Elektrokymogram; *des Gehirns*: Elektroenzephalogram, [...]  
(22) LYSA: utstråla lys, [...], (om lampa) brinna; [...] (om sol i. måne, ibl.) vara uppe; [...]

og med angivelse af brugsområde stilmæssigt, særsprogligt og regionalt (angivet med etiketter som barnesprog, nedsættende, vulgært, i sømandssprog, politisk, dialektalt, etc.) og af syntaktiske og kollokationelle begrænsninger (angivet med grammatiske termer, fx transitiv, eller med dele af kontekst) eller med rent relative angivelser af partiel synonymi eller hyponymi (angivet med etiketter som iblandt, specielt etc.), fx:

- (23) ADELSMAND: aristokrat, herremand, ædling; (spec.) baron, baronet, bojar, effendi, friherre, fyrste\*, grande, greve, [...]

*Webster's Collegiate Thesaurus* giver en relativ markering (med kapitælskrift) af, hvilket ord i en gruppe der er hovedopslagsord og dermed det mest generelle ord. Endvidere markeres vulgær brug og (med tegnet ||), at anvendelsen af et ord "may involve a problem of diction too complex for presentation in a thesaurus, or a restriction in usage" (s. 26a; det sidste kriterium hentyder til stigmatiserede og regionale former). Fx:

- (24) COTTAGE: HUT, || box, || caboose, camp, cot, lodge, shack, shanty

Hvis information om differentieret interrelatering inden for en gruppe skal bibeholdes for alle opslag, må gruppen placeres som samlegruppe under et umarkeret ord, eller også må opslagsord også kunne tildeles en markør. Det sidste er praksis i *Ord för ord* og i *Webster's Collegiate Thesaurus*, fx:

- (25) BRASSA PÅ: [...]; (vard. för) ta(ga) i, klämma i, (vard.) 'köra hårt'

- (26) || CABOOSE: HUT, || box, camp, cot, cottage, lodge, shack, shanty

Verbaliserede kommentarer som i eksemplerne (21-22) kan dække mange typer information og kan være åbne med hensyn til fx betydningsbeskrivelse. *Duden Sinn- und sachverwandte Wörter* skelner fx mellem fjorten nærmere definerede stilmarkører (dichterisch, bildungssprachlich, gehoben, umgangssprachlich, familiär, Jargon, salopp, derb, vulgär, abwertend, scherzhaft, emotionel, landschaftlich, verhüllend) og beskriver frit forskelle i betydning.

Brug af en sådan mere detaljeret etikettering kan imidlertid ikke ses som en systematisk angivelse af relationelle forhold. Det må være begrænset, hvor mange områder der kan dækkes af redaktionelt definerede markører, og markering herudover må være styret af, hvad der kan formuleres kort, sammenlign fx behandlingen af kohyponymer under *ELEKTROGRAMM* og *LÖFFEL* i eksemplerne (21) og (11) ovenfor. For ordbogsbrugeren kan tilstedeværelsen en sådan sporadisk markering tilsløre, at der også er relationelle forskelle mellem ord, som ikke er markerede.

En mere gennemgribende markering af relationerne i en ordgruppe forudsætter, at disse markeres, ikke ved absolutte beskrivelser, men som relative inden for gruppen, dvs. ved overordnede relationstyper som strikt synonymi, partiel synonymi, kohyponymi, hyperonymi og hyponymi.<sup>5</sup>

Ved relativt markeret differentiering inden for en gruppe må markeringerne være relateret til et fællespunkt i form af en definition. Markeringerne kan så gælde i forhold til gruppen som helhed, dvs. i forhold til det, der definerer denne (fx i form af en definition under et hovedopslagsord som i *Webster's Collegiate Thesaurus*). Eller markeringerne kan gælde i forhold til de enkelte opslagsord. I så fald vil markeringerne variere, jf. (8).

Det vil imidlertid ikke være muligt at angive relationer mellem alle medlemmerne i en gruppe. Man kan angive identitet med fællespunktet, og man kan angive forskellige afvigelser herfra (fx med hensyn til distribution og relationer som kohyponymi, hyperonymi og hyponymi). Når der er tale om identitet, er de ord, der markeres herfor, også indbyrdes identiske. Men når der er tale om ord, der markeres som afvigende fra fællespunktet, kan de indbyrdes relationer mellem disse ikke angives (med mindre man bruger meget komplicerede markeringssystemer). Hvis = og || står for henholdsvis identitet og ikke identitet, vil relationerne i:

(27) =a, =b, ||c, ||d

kunne læses sådan, at *a* og *b* er identiske, og *c* og *d* forskellige fra *a* og *b*, men man vil ikke kunne aflæse noget om *c* og *d*'s indbyrdes forhold.

En relativ markering må bl.a. derfor, og i højere grad end en absolut markering, ses i sammenhæng med en betydningsordbog.

### Hvorfor synonymer i betydningsordbøger?

Integration af makrostrukturelle semantiske relationer i betydningsordbøger, herunder eksplicit angivelse af synonymer, har været behandlet især i tysk leksikografisk litteratur.

Som de funktioner, systematiske synonymangivelser kan have i en betydningsordbog, anfører Herberg (1989) s. 154:

- bidrag til betydningsforklaring (svarende til den traditionelle brug af synonymer til parafrase af lemma),
- synliggørelse af en semantisk systematik i form af synonymforbindelser i lemmabestanden,
- information om ordnede alternative udtryksmuligheder ved tekstproduktion.

De to første informationsfunktioner har betydning for den, der bruger ordbogen som definitionsordbog. Synliggørelsen af betydningsmæssige strukturer i ordbogen kan have den funktion, at et opslag placeres i et semantisk rum, hvorindenfor opslaget betydning og brug yderligere kan afgrænses ved andre opslag.

Den tredje funktion gælder for den, der anvender synonymer til at variere og præcisere sit sproglige udtryk. Synonymbrugereren søger ikke blot stilistisk variation på udtryksplanet i form af fuldstændig synonymi (som fx *karbonade/krebinet*, *ringdue/skovdue*), men nok i højere grad syntaktisk og kollokationelt bundne varianter og udtryk for nuancer i betydning og ekspressivitet.

For synonymbrugereren er det en fordel, at definitioner, eksempler og konstruktioner er redigeret sammen med synonymer og befinder sig på samme sted, således at man mere umiddelbart kan gå til den sproglige information, som brugen af listede synonymer kræver (jf. som omtalt ovenfor opfordringen i nogle synonymordbøger til brugereren om at støtte sig til en betydningsordbog).

Hvis disse forskellige funktioner skal tilgodeses, skal der i betydningsordbogen inkluderes synonymer baseret på et bredere synonymibegreb, og dermed et større antal end i fx *Collins Cobuild English Language Dictionary* og *Svensk Ordbok*. Og der skal være en skelnen mellem strikt synonymi og forskellige andre former for betydningsbeslægtethed (både af hensyn til definitionsbrugereren og for at lette synonymbrugerens søgning).

Differentiering inden for de enkelte synonymgrupper sker bedst i form af relative relationelle markører som omtalt ovenfor. Disse markører (eller eventuelt fravær af disse) henviser så til information under de pågældende opslagsord. En sådan integration af synonymer i en betydningsordbog stiller hermed ideelt set også krav om, at der ved redigering af betydninger, brugsoplysninger og eksempler tages hensyn til synonymer, så de enkelte artikler belyser de forskelle, som differentieringsmarkørerne er signal om.

## Noter

1. Synonym og synonymi står her (sådan som det er tilfældet i titlerne på flere ordbøger) som en bred betegnelse, der også dækker hyponymirelationer og fjernere betydningslægtsskab. Beskrivelsen i artiklen omfatter da også ordbøger, der som fx *Duden Die sinn- und sachverwandten Wörter* har karakter af begrebsordbøger. Hvor det er nødvendigt benyttes betegnelsen betydningsbeslægtede ord over for snæver eller strikt synonymi.

2. I dette og de følgende citater skrives opslagsord med store bogstaver, og henvisning markeres med \*, - uanset hvordan præsentationen heraf er i det citerede værk. I eksemplerne (11) og (21) er Dudens gruppeadskiller, en eller flere prikker, erstattet med semikolon.

3. Jf. Schippan (1987) s. 225: "Synonymgruppen haben den Charakter von Teilsystemen und sind damit im Sinne der Prager Schule durch Zentrum und Peripherie gekennzeichnet. Sie umfassen alle die Synonyme, die sich durch Nebenmerkmale (untergeordnete Merkmale) vom Grund- oder Leitsynonym, vom Archilexem, abheben und mit ihm gemeinsam in der Mehrzahl der Kontexte auftreten können."

4. Blumenthal/Lemnitzer/Storrer (1987) specificerer elementerne ved henvisning som "Verweisursprung" (E), "Verweisziel" (A) og "Verweisangabe" (a\*) bestående af "Verweiszielrepräsentant" (a) og "Verweissymbol" (\*). Inden for denne terminologiske ramme er henvisningsfunktionen i (10) yderligere baseret på en "Verweisursprungsrepräsentant" (e).

5. Sådanne relationstyper kunne, på grundlag af fælles betydningskerne, være defineret ud fra implikationsmæssige og distributionelle kriterier (hvor distribution står for tekstlige restriktioner i bred forstand, fra særsproglige til syntaktiske). Der kunne så være tale om markering af relationer som:

- Gensidig implikation, dels i forbindelse med distributionel identitet, dvs. substitution mulig i alle kontekster, og dels i forbindelse med ikke distributionel identitet, dvs. substitution ikke mulig i alle kontekster. Det sidste ville kunne differentieres videre i inkluderende, overlappende eller komplementær distribution og i forskelle med hensyn til særsproglig, stilistisk eller syntaktisk distribution.

- Ensidig implikation, dels i form af hyperonymi, dvs. ensidig substitution er mulig (men med forskydning i betydningspotentiale), dels i form af hyponymi, dvs. substitution er ikke mulig.

- Ingen implikation (men stadig fælles betydningskerne): kohyponymi, dvs. substitution er ikke mulig.

## Litteratur

Blumenthal, Andreas & Lothar Lemnitzer & Angelika Storrer. 1988. Was ist eigentlich ein Verweis? Konzeptionelle Datenmodellierung als Voraussetzung computergestützter Verweisbehandlung. I: Gisela Harras, udg. *Das Wörterbuch. Artikel und Verweisstrukturen*: s. 351-373. Düsseldorf.

*Collins Cobuild English Language Dictionary*. 1987. John Sinclair mfl. London/Glasgow.

*Collins Concise Dictionary and Thesaurus*. 1991. Glasgow.

*Dansk Synonymik*. 1872. P.E. Müller. 3. udg. ved F.P.J. Dahl. København.

*Dansk synonymordbog*. 1986. Ulla Albeck mfl. 7. udg. ved Bent L. Møller. København.

- Duden Bedeutungswörterbuch*. 1985. Wolfgang Müller mfl. 2. udg. Mannheim/Wien/Zürich.
- Duden Sinn- und sachverwandte Wörter*. 1986. Wolfgang Müller. 2. udg. Mannheim/Wien/Zürich.
- Herberg, Dieter. 1989. Synonymische Beziehungen im Wortschatz und Wörterbucheinträge. I: *Hermes 3-1989*: s. 143-160. Århus.
- Longman Dictionary of the English Language*. 1984. Harlow, Essex.
- Ordbog over det danske Sprog*. 1918-56. København.
- Ord för ord*. 1977. Herbert Friedländer & Johan Palmér. 2. udg. ved Herbert Friedländer. Stockholm.
- Politikens Nudansk Ordbog*. 1990. 14. udg. København.
- Politikens Synonymordbog*. 1991. Allan Karker. 10. udg. København.
- Schippan, Thea. 1987. *Lexikologie der deutschen Gegenwartssprache*. 2. udg. Leipzig.
- Svensk ordbok*. 1986. Sture Allén mfl. Stockholm.
- The Oxford Thesaurus*. 1991. Laurence Urdang. Oxford.
- The Penguin Modern Guide to Synonyms and Related Words*. 1971. S.I. Hayakawa. Revideret udg. ved P.J. Fletcher. Harmondsworth, Middlesex.
- Webster's Collegiate Thesaurus*. 1988. Springfield, Mass.
- Webster's Dictionary of Synonyms*. 1951. Springfield, Mass.

Anne-Line Graedler

## Orientering om prosjektet "Norsk anglisismeordbok"

Ordboksprosjektet er en del av et større prosjekt som går ut på å kartlegge bruk av engelsk i norsk, og engelsk innflytelse på norsk språk. Ordboka skal dokumentere engelsk innflytelse på ordplanet, samtidig som den skal kunne brukes som en type fremmedordbok. Artikkelen tar opp sider ved ordbokas funksjon, materialet den er bygd på, og hvilke kriterier som ligger til grunn for utvalget av ord. Siste del av artikkelen presenterer utforminga av boka på mikrostruktur-nivå, med illustrasjoner i form av tenkte ferdige ordboksartikler.

### 1. Innledning: Prosjektet "Engelsk i norsk språk og samfunn"

Ordboksprosjektet er en del av det større forskningsprosjektet "Engelsk i norsk språk og samfunn", som har vært i gang ved Institutt for britiske og amerikanske studier, Universitetet i Oslo (UiO), siden januar 1990.

En viktig del av arbeidet går ut på å samle og registrere i en database engelsk og engelsk-påvirket ordmateriale som brukes i norsk. Denne databasen, som i dag består av rundt 13 000 poster (belegg), danner grunnlag for ordboka, og også for en monografi innenfor området.

### 2. Bakgrunn

Etter annen verdenskrig har vi opplevd at en stadig økende mengde engelske ord og uttrykk på forskjellige måter blir tatt opp i norsk. I forordet til *Nyord i norsk 1945-1975* regnes det med at 80-90% av alle nyord med utenlandsk opphav på en eller annen måte stammer fra engelsk. Fenomenet opptar ikke bare forskere, men i høy grad også vanlige nordmenn og -kvinner, og mye av oppmerksomheten har form av bekymring på vegne av norsk språk. I løpet av det siste året har det vært gjennomført i hvert fall to større kampanjer med bekjempelse av "anglonorsken" som hovedmål: Norsk språkråds "Aksjon for språklig miljøvern", og Noregs Mållags kampanje under slagordet "Elsk NORSK - ikkje engELSK!". Bakgrunnen for ordboka er et ønske om å samle og dokumentere informasjon om dette fenomenet med utgangspunkt i ordforrådet, og samtidig lage ei bok som skal kunne brukes og være til nytte. Det kunne innvendes mot boka at det blir kunstig å skille mellom engelske lånord og innlån fra andre språk. Til forsvar for en slik begrensning kan man si at ikke bare er lån fra engelsk totalt dominerende, samlet representerer de også en del spesielle problemer når det gjelder tilpasningen til norsk, f.eks. i forbindelse med uttale og bøyning. Det å samle engelske lånord i ei ordbok er for øvrig heller ingen ny tanke (se litteraturliste).

### 3. Materialet

Materialet i databasen som skal ligge til grunn for ordboka, består av skriftspråkseksempler fra dagsaviser, tidsskrifter og ukeblad, og også noe fra skjønnlitterære verk. I tillegg har vi fått benytte et større materiale fra nyordsamlinga til Avdeling for leksikografi, UiO. Dessuten har vi noe midler til å søke i databasen til Aftenpostens Elektroniske Informasjonstjenester, som vi bl.a. kan bruke til å komplettere eksempelsamlinga. I arbeidet med ordboka konsulteres også andre norske ordbøker (se litteraturliste).

Registreringa av materialet skjer ved hjelp av et databaseprogram, askSam. Hver post inneholder ett eksempel på bruk av et engelsk ord eller uttrykk i norsk med kontekst og kildehenvisning, pluss felter med informasjon om den opprinnelige engelske formen, typen av lånord, grammatisk informasjon, syntaktisk funksjon, i hvilken grad ordet er formelt integrert i norsk, om ordet er markert som fremmedord, grafisk eller språklig, om referenten er norsk, engelsk osv., ev. betydningsavvik, emneområde og teksttype.

### 4. Ordbokas funksjon

Sentralt i forbindelse med utforminga av ordboka står spørsmålet om hvilken funksjon den er tenkt å ha. Er det behov for den type informasjon ordboka vil gi? Hvem er i så fall brukerne? Hva vil de bruke boka til, og hvorfor?

På mange måter faller ei bok av denne typen mellom flere stoler: Den er kanskje nærest beslektet med fremmedord- og nyordbøker, men har også flere likhetstrekk med tospråklige ordbøker. I utgangspunktet er boka tenkt som ei passiv ordbok for norske brukere, som vi antar vil ha interesse også utenfor akademiske sirkler. De potensielle brukerne vil kunne være språkfolk og andre spesielt interesserte, for dem vil hovedinteressen kanskje ligge i bokas funksjon som dokumentasjon. En annen gruppe er de som i tillegg ønsker å bruke boka som fremmedordbok, og muligens som kilde til avløserord for en del av lånorda. Dette vil innebære at boka også vil kunne ha en aktiv funksjon. En tredje gruppe er f.eks. (engelsk)lærere, som vi håper vil kunne bruke boka for å finne informasjon om viktige forskjeller i engelsk kontra norsk ordbruk og -betydning.

Boka er tenkt å være primært deskriptiv, og det vil derfor bli viet forholdsvis mye plass til eksempel materiale med kildereferanser.

### 5. Utvalg av ord

Termen *anglisisme* kan referere til både såkalte direkte og indirekte lånord. Direkte lån kalles ord som er lånt i den form de har på engelsk, og deretter tilpasset det norske systemet i større eller mindre grad, f.eks. *boks*, *tøff* og *jogge*. Det kan også være komponenter i sammensatte ord, f.eks. *offroad-sykkel* og *kles-minded*, og vi kan også ta med de såkalte psevdolåna, dvs. ord dannet av engelsk ordmateriale, men hvor den engelske originalen mangler; dette er som oftest sammensatte ord, f.eks. *softice*. Indirekte lån kan vi kalle engelsk-påvirkede ord hvor selve ordformen allerede fins i norsk, eller er dannet av norsk ordmateriale. Det kan være betydningslån, dvs. norske ord som får en ny eller utvidet betydning som følge av påvirkning fra en tilsvarende form på engelsk, f.eks. *karakter* 'tegn' og *album* 'grammofonplate', eller oversettingslån, som er sammensatte ord som oversettes ledd for ledd til norsk, f.eks. *frynsegode* og *kroppbygger*.



Ut fra tanken om at boka skal være en dokumentasjon av engelsk innflytelse på norsk på ordplanet, burde vi vel ta med alle typer lån som er nevnt ovenfor. Det er imidlertid en del usikkerhet knyttet til mange av de indirekte lånorda: Enkelte av orda som kunne kalles betydningslån, resulterer i bare svært subtile betydningsendringer, f.eks. når *ballade* utvides til også å gjelde populærmusikk. Dessuten er en del av det man kunne kalle oversettingslån, ikke ordrett oversatt, f.eks. *hjertestimulator* for *pacemaker* og *rullebrett* for *skateboard*. De fleste av oversettingslåna vil også være greie å forstå, så det er ingen grunn til å ta dem med som fremmedord. Direkte lån og mange av betydningslåna er derimot ofte vanskelig å forstå uten kjennskap til engelsk. Det er derfor disse typene vi vil konsentrere oss om, med en naturlig hovedvekt på direktelån.

I tillegg til enkeltord vil det også være aktuelt å ta med fraser og lengre uttrykk, f.eks. *country & western*, *flippe ut*, *fed up*, *the show must go on*, *hands up*. En del hel- og halvoversatte uttrykk, f.eks. *blåse opp* 'forstørre', *levere varen* og *stjele showet* vil også bli vurdert.

I utvalg av ord vil vi også ta hensyn til lånordas alder. Det er som regel umulig å vite eksakt når et ord er blitt lånt inn i språket, og enda vanskeligere å angi tidspunktet når det opphører å være et lånord. Vi vil ikke ta med ord som er kommet inn tidligere enn ca. midt på 1800-tallet, og på grunn av materialets sammensetning vil det være en naturlig vekt på lånord som har vært i bruk i etterkrigstida, og særlig fra 1960 og framover. Frekvens og spredning er en annen relevant faktor: Orda som tas med, må forekomme minst to ganger i to ulike tekster av forskjellige forfattere. Videre vil vi konsentrere oss om ord fra allmennspråket, det vil si som forekommer i publikasjoner beregnet på et bredt publikum av ikke-spesialister, f.eks. i dagsaviser. Ord som alltid forekommer i anførselstegn, eller som alltid blir forklart eller oversatt med norske ord, vil vi vurdere å utelate. På den andre sida er det ikke nødvendigvis så interessant verken å dokumentere bruken eller forklare betydningen av gamle, veletablerte og/eller hyppig forekommende ord som f.eks. *sjal*, *film*, *sport*, *stessa*, *kutte*, *tiping* - ord som de fleste språkbrukere sannsynligvis tenker på som norske. En mulighet er å gi slike ord en type "minimal-oppslag", slik at de kommer med i samlingen, men ikke tar opp unødig plass. Her, som ellers, gjelder selvsagt at det må brukes skjønn.

## 6. Kriterier for oppstilling av ordboksartikler

En foreløpig planlagt størrelse på ordboka er ca. 5 000 oppslag. (Til sammenlikning kan nevnes at f.eks. *Bokmålsordboka* har anslagsvis 2 200 ord som på en eller annen måte stammer fra engelsk.) Boka skal ordnes alfabetisk. I det som følger vil jeg gjøre rede for de prinsipper vi følger for organiseringa av selve ordboksartikkelen, eller ordbokas mikrostruktur.

Ordboksartiklene skrives inn som poster i en database, og vi har fått hjelp til utforminga av Håvard Hjulstad fra Rådet for teknisk terminologi. Eks. 1 viser en slik post. I det følgende vil jeg også referere til eksemplene i appendikset på slutten av artikkelen, som for lesbarhetens skyld forsøksvis er satt opp som foreløpig "ferdige" artikler.

1 O shop

U

G subst. m. <UT -(p)en>, flt. <UT -(p)er>, <UT -(p)ene>

D <BET <K ofte i sammensetn.> butikk, særl. (liten) spesialbutikk> <ORDB BOB, NOB> <BEL1 <TEXT Heller ikke er det noen strid om at Interessekontoret kan opprette en liten shop> <KIL Dagbl. 26.11.60, 15>> <BEL2 <TEXT Bacchus vinshop er en liten perle midt i Skagerak City.> <KIL Aftp. 10.2.89, 46>>

SMS coffee-, data-, intim-, ski-, sommer-, souvenir-, tax-free-, vin-, <X work->  
 REF sjappe  
 =

Artikkelen vil bestå av et oppslagsord (feltet merket O i eks. 1), dvs. ordets grunnform og offisielt godkjente rettskrivningsnorm, dersom ordet er normert. Hvis ordet ikke er normert (se eks. 8 *showbiz*), eller flere enn en form er godkjent norm, vil vi bruke den mest frekvente formen i vårt materiale som oppslagsord. Homografer skal skilles både ut fra formelle og semantiske kriterier: Like former som hører til forskjellige ordklasser vil bli ført opp separat (se eks. 6 og 7 *show*), det samme gjelder ord med svært forskjellig betydning på norsk, f.eks. en del "dobbelte lån" (ord som antas å være lånt inn to ganger med forskjellig betydning). Vi opererer ut fra prinsippet "hvert ord sin artikkel", men sammensetninger vil få egne artikler bare der de er leksikaliserte (se eks. 6 *show* med kryssreferanser til f.eks. *showbiz*, *showbusiness* og *showdans*).

Vi vil også oppgi alternative skrivemåter som vi har dokumentert i materialet vårt, ordnet etter frekvens (se eks. 8 *showbiz*).

Etter grunnformen(e) ønsker vi å angi uttale av lånordet i norsk (det tomme feltet merket U i eks. 1). Dette er en problematisk del av arbeidet, så lenge vi også ønsker å være deskriptive, og dessuten baserer boka på et skriftlig materiale. Uttale er delvis angitt i en del norske ordbøker, og en mulighet ville være å overta denne for de aktuelle orda. Der vi ikke har ordbøker å støtte oss til, vil introspeksjon og informanttesting, eller en kombinasjon, være de aktuelle mulighetene; dette er selvfølgelig i stor grad et spørsmål om ressurser. I forbindelse med uttaleangivelse dukker også normative aspekter opp: Skal man velge en fornorsket uttale og hvilke(n) uttalevariant(er) skal den i så fall være basert på, eller skal man velge en mest mulig korrekt engelsk uttale, og skal denne i så fall være basert på britisk uttale, amerikansk uttale, eller annet. Notasjonssystemet vil også måtte vurderes, primært ut fra hensynet til de potensielle brukerne av boka.

Deretter følger grammatisk informasjon (feltet merket G i eks. 1). Her blir det angitt ordklasse, ev. kjønn eller bøyingsklasse, og de bøyingsendelser som er dokumentert i vårt materiale. Dette har både ulemper og fordeler: På den ene sida får man lite grammatisk informasjon om lite frekvente ord (se eks. 12 *singback*), og antall bøyingsendelser vil variere svært fra ord til ord; enkelte ord kan også ha alternative bøyingsmønstre (se eks. 6 *show* og eks. 14 *sjanti*). På den andre sida unngår man å tvinge alle former inn i ett rammeparadigme, som kanskje stemmer dårlig med den faktiske bruken av ordet (se eks. 9 *showbusiness*, som bare forekommer i entall). Mengden og variasjonen i bøyingsendelser vil også indirekte gi informasjon om i hvilken grad ordet er formelt eller morfologisk integrert i det norske systemet. Der vårt materiale er mangelfullt, vil det være aktuelt å supplere med den informasjonen som fins i ordbøker.

I de tilfeller hvor lånordets form i norsk avviker fra den engelske opprinnelige formen, eller hvis det fins flere enn en form, vil den engelske formen bli angitt i parentes (se eks. 3 *shoppe*). Dette feltet benyttes også til andre typer etymologisk informasjon, f.eks. opplysninger om ordets (kjerne-)betydning på engelsk (se eks. 12 *singback*).

Deretter følger en definisjon (feltet merket D i eks. 1). Noen ganger kan det være aktuelt å begynne denne med kommentarer vedrørende stilnivå og ev. fagområde. Selve definisjonen (feltet merket BET i eks. 1) skal være ei kort og i størst mulig grad enkel forklaring av ordets betydning, og kan bestå av synonymymer og/eller flere ord. Synonymene

kan også være såkalte avløserord, men det kommer ikke til å være noe poeng å anbefale disse framfor de engelske lånorda (slik det er gjort i f.eks. *Nyord i norsk 1945-1975*).

Resten av informasjonen i artikkelen vil kanskje være mest interessant for dem som er spesielt opptatt av engelske lånord: Etter forklaringa følger en angivelse av ordets forekomst i en del norske ordbøker (feltet merket ORDB i eks. 1; forkortelsene er forklart i litteraturlista), med årstall for første forekomst der det er angitt. Dette vil, foruten at det fungerer som en henvisning, også si noe om ordets alder og utbredelse, og den sosiale integrering ordet har i norsk.

Så følger eksempler på ordet brukt i norsk (feltet merket BEL i eks. 1). Beleggene består av en kort tekst (feltet merket TEXT i eks. 1), med tilhørende kildehenvisning (feltet merket KIL i eks. 1), og de vil bli valgt ut fra følgende kriterier: De skal best mulig illustrere bruken av det aktuelle ordet, det eldste belegget bør tas med, alle betydningsnyanser skal dekkes, og det er en fordel om betydningen også helt eller delvis framgår av eksemplet og sammenhengen ordet er brukt i; forskjellige bøyingsmåter eller skrivemåter kan gjerne tas med. Eksempelene stilles opp i kronologisk rekkefølge.

Etter eksemplene kan det komme kommentarer (feltet merket KOM i eks. 1) knyttet til f.eks. spesiell bruk av ordet på norsk i forhold til engelsk (se eks. 7 *show*).

Til slutt i artikkelen kommer en del tilleggsdokumentasjon; fraser eller (faste) uttrykk som ordet inngår i, norske avledningsformer, og sammensetninger der ordet er en komponent, blir angitt uten kommentarer, men med krysshenvisning (i form av en pil i eksemplene) der en av disse har eget oppslag (se eks. 6 *show*).

Programmet gir dessuten mulighet for å legge inn krysshenvisninger og kommentarer hvor som helst gjennom hele artikkelen.

## 7. Problemer/oppsummering

Arbeidet er enda i en tidlig fase, og det kan fortsatt være aktuelt å foreta visse justeringer. Den todelte funksjonen som ordboka er tenkt å ha, er den største generelle vanskeligheten, og det som medfører de største ordbokstekniske problemene: På den ene sida ønsker vi at boka skal være en dokumentasjon av språkbruk med utgangspunkt i et (relativt) sluttet tekstmateriale, og at den derfor skal være deskriptiv. Dette oppnår vi ved å oppgi alle former, eller noen ganger bare de formene vi har dokumentert i vårt materiale, og eksempler fra faktisk språkbruk. På den andre sida vil vi at boka skal kunne tjene praktiske formål, og brukes av ikke-spesialister. Vi renonserer derfor på det deskriptive idealet, og angir uttale av ord, og normerte former foran mer frekvente former. Men en slik kompromissløsning er sannsynligvis ikke til å unngå dersom ei bok av denne typen skal kunne utgis i Norge.

I tillegg dukker det selvsagt stadig opp detaljspørsmål i forbindelse med redigeringsarbeidet. En del problemer kunne sannsynligvis vært unngått hvis tilpasningen av de to databaseprogrammene vi bruker, hadde vært enda mer gjennomtenkt. Det å ha et såpass stort eksempelmateriale lett tilgjengelig på PC-en når vi skriver, er likevel av stor verdi.

**Appendiks: Eksempler på ordboksartikler**

(Merk at i uttaleangivelsen er "ng-lyden" representert ved bokstavsekvensen -ng-.)

- 2 **shop** /ʃɑp/ subst. m. *-(p)en, flt. -(p)er, -(p)ene* (ofte i sammensetn.) butikk, særl. (liten) spesialbutikk (NRO 1931, BOB, NOB, TAN) *Heller ikke er det noen strid om at Interessekontoret kan opprette en liten shop (Dagbl. 26.11.60, 15) Bacchus vinshop er en liten perle midt i Skagerak City. (Aftp. 10.2.89, 46)*  
 sms coffee-, data-, intim-, ski-, sommer-, souvenir-, tax-free-, vin-, →workse også →sjappe
- 3 **shoppe** /ʃɑpe/ v1 (eng. *shop*) handle, gå i butikker (BOB, NOB, TAN) *Så de ikke behøver kle seg om når de skal ned i selfservicen og shoppe. (Dagbl. 15.6.60, 8) For henne kan det bli en opplevelse. Tid til å shoppe om dagen og kanskje tar dere en svingom om kvelden (Farmand 1974:11, 6) avl shopper*
- 4 **shopping** /ʃɑping/ subst. m. *-en* handling, handletur, det å gå i butikker (BOB, NOB, TAN) *Om dagen når man farer rundt i biler og på shopping. (Dagbl. 22.1.60, 9) Det ser ut som han har vært på shopping for Nansen før Grønlandsferden (Michelet MPM, 293)*  
 sms jule-, tele-, vindus-, window-, mulighet, →senter, -sentrum, -sted, -tur, -vindu, -virksomhet
- 5 **shoppingsenter** shopping-senter shopping-centre /ʃɑpingsenter/ subst. n. *-et el. -tret, flt. -tre el. -tra, -trene* (eng. *shopping centre, US center*) kjøpesenter, mange butikker samlet i ett lokale (BOB, NOB) ... *med stoppested like ved shoppingsentret. (Dagbl. 19.11.60, 12) Et annet utviklingstrekk er at de store shoppingsentrene har mer opplevelse og flere aktiviteter å tilby kundene. (Aftp. 1.3.89, 20)*
- 6 **I show** /ʃɑv/ el. /ʃɑo/ subst. n. *-et, flt. -, -s (-s er sjelden), -ene* 1 forestilling, oppvisning, revy; oppstyr, ståhei (NRO 1939, BOB, NOB, TAN) *showet kan denne gang by på et virkelig bombenedslag av humørfyllt sensualitet og sex-betont gjøglerglede (Dagbl. 1960:177, 2) Som en fugl fønix skal den nye norske banken reise seg fra asken ... Showet som skal danne grunnlaget for en ny image, koster femten millioner kroner.) (Aftp. 11.4.90, 5) 2 (særl. i sammensetn.) presentasjon og utstilling av (et større utvalg) varer, ofte med spesielle tilbud *Årets store støvletshow foregår nu (Mgbl. 1963:296, 8) Indre Østfolds største frukt-/grøntavd. Lørdag 18. DOLE Banan-show (Obs!-reklameavis 16.2.89, 2 ann.) 3 (særl. i sammensetn.) sjå, oppvisningskonkurranse fra →best in show, →stjele showet, →the show must go on, time will show avl →showe*  
 sms banan-, Broadway-, Casino-, crazy-, →drag-, enkvinne(s)-, enmanns-, Euro-, fashion-, flo-, floor-, freak-, fryse-, game-, gate-, Greatest Hits-, helaftens-, hunde-, is-, jubel-, jule-, kjempe-, kunstløp-, laser-, lys-, lørdags-, mega-, milkshake-, misse-, mote-, motor-, multi-, multi-media-, musikk-, nonstop-, →one-man-, →one-woman-, pop-, rock(e)-, scene-, shanty-, skate-, snakke-, Spellemann-, spesial-, stor-, strip together-, striptease-, støvlett-, super-, →talk-, topp-, toppless-, trailer-, trend-, trylle-, tv-, utstys-, vare-, -aktør, -artist, -betont, →biz-, -boat, →business, -båt, -cruise, →dans-, -fag, -festival, -film, -form, -girl-, kapellmester, -keeper, -kis, -konkurranse, -*

man, -mann, -messig, -minne, -musikk, -nummer, -preg, -preget, -produkt, -program, -prosess, -rom, -room, -scene, -sensasjon, -stil, -time

7 II **show** /ʃåv/ adj. morsom, artig *sorte, showe lakksko m/gullpynt og kraftig hæl (Morgenbl. 1969:271, 11 ann.) Det er veldig mange showe politikere, men det er når de er utenfor politikken. (Dagbl. 3.8.91:2, 2)*

**kom** på eng. brukes ikke *show* som adj., denne bruken kan være påvirket av da. *sjov* 'morsom'

8 **showbiz** show-biz show-bizz show-biss /ʃåvbis/ el. /ʃåo-/ subst. m. (eng. *showbiz*) kortform for →*showbusiness* *F tror folk er lei av tingel-tangel og lettvent showbizz. (Dagbl. 9.6.90:2, 21) Grand old man i norsk show-biz, Leif Juster, var ikke mindre sjarmerende (Dagbl. 29.8.90, 37)*

sms -tidsskrift, -verden

9 **showbusiness** show-business show business /ʃåvbisnis/ el. /ʃåo-/ subst. m. -en (eng. *show(-)business*) underholdningsindustri, kommersiell drift av teater, revy, o.l. (BOB, NOB) *denne 34-årige entertaineren, som hører til de absolutte toppene i amerikansk showbusiness (Dagbl. 7.10.60, 16) Jeg synes folk i showbusiness er noen store drittsekker, og det er altfor få som tør å si dette. (Dagbl. 28.04.90:2, 11)*

se også →*showbiz*

10 **showdans** /ʃåv-/ el. /ʃåo-/ subst. m. -en, flt. -er, -ene seque, øvelse i sportsdans hvor det danses nokså fritt til latin-amerikanske rytmer *I showdans eller "seque", som er den internasjonale betegnelsen, er det bare ett par på gulvet ad gangen, og de danser etter latin-amerikanske rytmer i fire og et halvt minutt. Svært likt kunstløp. (A-mag. 2.3.91, 63)*

sms -VM

11 **showe** /ʃåve/ v1 (av →*show* I) gjøgle, lage show, ha det moro (BOB, NOB, TAN) *DFI spiller, scorer, shower og jubler så det holder. (Østl.Bl. 27.8.90, 9) Vi kommer ikke til å "showe" slik som dehar gjort, men servere et enklere konsept, uten sketsjer, kostymer og uteopptak. (Dagbl. 11.1.91:2, 5)*

12 **singback** /ʃingbæk/ subst. (eng. *sing* 'syng' + *back* 'tilbake') opptreden hvor en artist synger til akkompagnement som er spilt inn på lydbånd *Singback, der artisten synger til musikk på bånd, er ... tillatt med visse begrensninger. (Dagbl. 5.5.90:2, 22)*

13 **single** /ʃingel/ subst. m. -n, flt. -r, -ne liten grammofonplate med en sang på hver side, avspilles på 45 rpm (NIN 1964, BOB, NOB, TAN) *Det finns plate med Robertino både på single og E.P. (Dagbl. 8.12.60, 6) I sin anmeldelse av singlen ... skriver Billboard ... (VG 29.3.89, 33)*

sms hit-, -hit, -plate, -pris

14 **sjanti** shanty /ʃanti/ subst. m. -en, flt. -er, -es (-es er sjelden), -ene (eng. *shanty*) sjømannssang, oppr. arbeidssang ombord på seilskute (NRO 1933, BOB, NOB, TAN) *Jeg har lyst*

å trekke fram Erik Byes egen shanty-vise. (Dagbl. 14.5.60, 1) TJ reiste seg opp og lot avsyngne en slags shanty om en person ved navn Theodor (Kap. 1990:20)

sms -gruppe, -show, -vise

15 **sjapp(e)** /ʃap(e)/ subst. f. -ppa el. m. -(e)n, fl. -(e)r, -(e)ne (formen *sjapp* og m. er sjelden) (eng. *shop*) liten butikk, bu el.lign. (NRO 1891, BOB, NOB, TAN) *Jeg valgte en såkalt sjappe - en slik som ser snusket ut, men lukter rent* (Dagbl. 1960:151, 9) *I den eksklusive sjappa i Sackville Street ...* (Michelet MPM, 31)

sms oldies-

se også →shop

## Litteratur

*Anglismeordbøker og -prosjekter:*

Alfaro, R.J. 1970. *Diccionario de anglicismos*. 3. utgave. Madrid.

Bonaffé, Édouard. 1920. *Dictionnaire étymologique et historique des anglicismes*. Paris.

Carstensen, Broder m.fl. Under arbeid. *Wörterbuch der Anglizismen im heutigen Deutsch*. Gesamthochschule Paderborn.

Filipović, Rudolf. 1990. Rječnik anglicizama u hrvatskom ili srpskom (Ordbok over anglismer i sebo-kroatisk). I: *Anglicizmi u hrvatskom ili srpskom jeziku: porijeklo - razvoj - značenje* (Anglismer i serbo-kroatisk: opphav - utvikling - betydning). Zagreb.

Filipović, Rudolf m.fl. Under arbeid. *The etymological dictionary of anglicisms in European languages*.

Haensch, Gunther & Reinhold Werner. 1978. Un nuevo diccionario de americanismos: proyecto de la universidad de Augsburg. I: *Thesaurus* 33:1.

Neske, Fritz & Ingeborg Neske. 1972. *dtv-Wörterbuch englischer und amerikanischer Ausdrücke in der deutschen Sprache*. 2. utgave. München.

Pepník, Jaroslav. 1982. *Slovník amerikanismu* (Ordbok over amerikansimer). Praha.

Rando, Gaetano. 1987. *Dizionario degli anglicismi dell' italiano postunitario*. Firenze.

Rey-Debove, Josette & Gilbert Gagnon. 1981. *Dictionnaire des anglicismes. Les mots anglais et américains en français*. Paris.

Seltén, Bo. 1987. *Svensk ordbok*. Acta Wexionensia, Ser. 3. Language & Literature 1. Växjö.

*Norske ordbøker:*

TAN = Fossestøl, Bernt & Marius Sandvei, red. 1989. *Tanums store rettskrivningsordbok*. 7. utg. Oslo.

NOB = Hovdenak, Marit m.fl., red. 1986. *Nynorskordboka*. Oslo.

NRO = Knudsen, Trygve & Alf Sommerfelt, red. 1937-57. *Norsk riksmålsordbok*. Oslo.

BOB = Landrø, Marit Ingebjørg & Boye Wangensteen, red. 1986. *Bokmålsordboka*. Bergen, Oslo, Stavanger, Tromsø.

NIN = *Nyord i norsk 1945-1975*. 1982. Norsk språkråd. Oslo.

Erik Simensen

## Om ein plan for ei mellomnorsk ordbok — ordbok over skriftmålet i Noreg 1350-1550

Norsk språk frå tida 1350-1550 ("mellomnorsk") er lite granska og ofte vanskeleg å skjønne pga. sterkt varierende ortografi, nyutviklingar og lån. Gammalnorsk ordboksverk ved Institutt for nordistikk og litteraturvitskap (Universitetet i Oslo) vil derfor lage ei ordbok over norsk skriftmål i denne perioden, basert på diplom, lovtekstar og jordebøker. Ein tenkjer seg ei stor, eittbands ordbok som dekkjer ordforrådet (utan å vere tesaurisk) og gjer greie for tilhøvet mellom grafem og fonem i kjeldene.

1 I eldre nordisk språkhistorie er især to periodar lite kjende: (1) umordisk og (2) overgangsfasen frå gammalnorsk til nyare norsk, den som vanlegvis blir kalla mellomnorsk. Samanstillinga av desse to er på ein måte ikkje heilt rettferdig; for dei umordiske kjeldene er få og stort sett overleverte i eit alfabet som byr på særskilde problem i seg sjølv, medan kjeldene flyt langt rikelegare for yngre språksteg. Og seinmellomalderen står i ei særstilling for så vidt som vi veit mindre om språket i denne perioden enn om språket nærmast før, dvs. "klassisk" gammalnorsk.

2 Språket i denne perioden går tradisjonelt under nemninga "mellomnorsk"; ein term som det har vist seg vanskeleg å gje eit presist innhald, delvis fordi den meir går på ein prosess enn ein språktilstand. "Mellomnorsk" er inga stabil(isert) språkform, ser det ut til. Ikkje er det gammalnorsk (slik vi kjenner den frå den "klassiske" forma den har på 1200- og 1300-talet), og heller ikkje nynorsk; men kanskje likevel, og trass alt, nærmare det første enn det siste, iallfall dersom ein berre deler norsk språkhistorie i to hovudavsnitt med midten av 1500-talet som skilje. På denne bakgrunnen har ein då også kalla denne overgangsfasen for "sein" eller "yngre" gammalnorsk, delvis parallelt til nemningsbruken for dansk ("yngre middeldansk" (Skautrup 1947)) og især svensk språkhistorie ("den yngre fornsvenskan" (Wessén 1968)). Spørsmålet om kva ein skal kalle språket i denne tida må vel inntil vidare stå ope, det er heller ikkje det viktigaste. Hovudsaka er å få kartlagt perioden så godt som kjeldene tillet, og til dette trengst det ei fullstendig oversyn over heile tilfanget av ord og ordformer i kjeldene.

Nokre viktige og typiske endringar som skjer i denne perioden, er dei følgjande:

- 1) [Ð] går over til /d/ eller /t/
- 2) Motsetnaden nominativ : akkusativ i substantivbøying fell bort
- 3) Adjektiv misser kasusbøying
- 4) Personbøying fell bort i verb

5) Ordforrådet får eit mykje sterkare innslag av lån, frå dansk, svensk og nedertysk

Meir allment kan ein seie at det går føre seg ein gradvis overgang frå ein syntetisk til ein svakt analytisk språktype.

**3.1.1** M. Hægstad, D.A. Seip, P.N. Grøtvedt, E. Pettersen og andre har skrive ein del om denne perioden, men noka større, samla framstilling ligg ikkje føre. Heller ikkje finst det noka ordbok som dekkjer heile perioden. Fritznors ordbok når stort sett til midten av 1400-talet, Cleasby/Vigfusson går lenger, men då mest for islandsk, slik at svært mykje av det norske materialet er udekt. Rett nok finst eit glossar (ord- og sakregister<sup>1</sup>) til lovtekstar for tida 1388-1447, og nokre lister over lånord i Hægstads og Grøtvedts avhandlingar, men elles ingenting.

**3.1.2** Det faglege behovet for ei kartlegging (inventarisering), altså omsynet til norsk språkhistorie, er såleis klart til stades<sup>2</sup>. Av dei som elles kan tenkjast å ha interesse av ei slik ordbok, er vel først og fremst historikarar og andre som arbeider med kjelder frå norsk seinmellomalder, ikkje minst lokalhistorikarane, i det heile dei som primært er ute etter å forstå tekstar frå denne tida. Kjeldene er ofte vanskeleg tilgjengelege, både leksikalsk og på annan måte (grafisk, fonologisk og morfologisk). Dels møter ein mange nye ord (lånord især) som ikkje finst i ordbøker over norrønt og nyare norsk, dels finn ein ord som alt fanst i norrønt (det dreiar seg truleg om så mykje som 80-90 % av ordforrådet), men som ein ikkje (eller iallfall ikkje utan vidare) kjenner att pga. skiftande og inkonsekvent rettskriving. - Ein kan i ein del tilfelle søkje hjelp i ordbøker over eldre dansk og svensk (Kalkar, Söderwall), nedertysk (Lübber, Lasch ofl.) og eldre nynorsk (Aasen, Ross, Norsk Ordbok), men her er ein då avhengig av større bibliotek, og fleire av desse ordbøkene er ikkje heilt lette å finne fram i for andre enn nordistar.

**3.2.1** Vi treng å fange opp den norsken som hittil har falle mellom fleire stolar. Døme finst m.a. i eit brev frå Ryfylke 1534 (DN XXI 801), om synfaring og taksering av ein gard<sup>3</sup>:

...wore tilkraffde...at mæthe hwsszenn och *vmfforbæthrenn* pa hesby...Siidhen Senisthe *hwsse maade* pa Samme gardt war, primo *Jordt kellerenn* och *borghestwffwenn* och loffstwffwenn met theris *tilbehøringh* ...Sammeledis mæthe wy...*Machrelle nodhen* ffore xxv marc bwgille...Ffwenne vy ther ffore at hann met gardt och grynndt och *hwllveedher* ther pa fforbæret haffwer pa x marc bugille...

Her er fleire ord som ikkje finst i ordbøkene våre over eldre språk, såleis *vmfforbæthrenn* 'forbetring' (istf. gno. ábúð), *hwsse maade* 'hustaksering', *Jordt kellerenn* 'jordkjellaren' (utgrave eller jorddekt rom), *borghestwffwenn* 'tenarstova', *tilbehøringh* 'tilbehør', *Machrelle nodhen* 'makrellnota' (der *Machrelle* avløyser gno. qlunn) og *hwllveedher* 'holveiter' (attlagde grøfter). Ein kan innvende at desse orda dels er lette å skjønne (t.d. *vmfforbæthrenn*, *tilbehøringh*), dels kjende frå nyare ordbøker. Men dei fleste lesarar må nok tenkje seg litt om før dei kan identifisere det siste ordet (som er ein norvagsisme i dansk rettskriving).



Eit meir tradisjonelt språk møter oss i eit brev frå Vest-Agder 1543 (DN VII 760, sitert etter F. Hødnebo, *Utvalg av norske diplomer 1350-1550*, Oslo/København/Stockholm 1966, nr. 85). Det gjeld ein gardhandel med tilhøyrande grenseoppgang:

Wy ... gjør *Vitherlig* for alle y thette vort obnne breff At vy neruerindis hoes vor paa Sthødle ..., hørdum ord oc saagum haandeabaand teira mannum som saa heida aff einu halffum, Gunnar thoreson, End aff adrum halffum, Thorbiørnn Thorbiørnsønn. Sælde for<sup>ne</sup> Gunnar aderneffnde Thørbiørnn samre for<sup>ne</sup> *øidegaard* sthødle meder eit vdgiorde som either eigestødulen ..., met luthum och Lundum, som thilligger oc ligget haffuer fraa fornoc oc nyom innen gaards och vthan iij hornstaffuar imillom frelst oc aakierelaust for huerium mandum, oc alle Lauglige aagangum, *Thisligeste bekiende* oc saa forneffnd Gunnar Thoreson, at hand haffde opboret aff for<sup>ne</sup> Thorbiørn thorbjørnson heyle pending oc halffua saa mange som y theris kiøb kom, ... thorbjørn thilspurde ..., huore rethe marckeskipte,... Ther til suarad gunnar honom och sagde, ... Strengbeint fraa Paddefield oc thil ein *steinhage* som hafuer verit y gamell thid, ... strengbeint offuer Sthødlehougen, som eenn *skiffestein* bleff nedkast, ... Der paa holt hand den dom, at hand wilde *skillige* vide same skiffte, imellom Stødle oc møgland, ... oc haffue dij samme mend lyst oc *bekiend* samme ord for os, oc thennom vor oc kunict om samme marckeskipte. Thill ydermeire *Stadfestelsse* henger vi...

Ordforrådet viser fleire nyutviklingar på heimleg grunn: *øidegaard*, *steinhage*, *skiffestein*, *skillige*; og nokre lån: *Vitherlig* (i ny tyding, istf. gno. kunnigt), *Thisligeste* (istf. gno. með sama hætti el. 1.), *bekiende*, *bekiend* (istf. gno. kennast viðr eller lýsa), *Stadfestelsse* (istf. gno. staðfestu).

Ein annan type eksempel har vi i NgL 2. R. 3, nr. 57 A (=DN VI 610), ein rapport frå ein granskingskommisjon på 10 menn, oppnemnde av kongen, i hovudsak rådmenn i Bergen, frå 1490:

... Hollendrene sto them myghet til skadhe ok *forfangh* offwer Strandh hær i Bærgwen. ... Mæn kjøpmannen hadhe priwilegiær som kwngth Cristiærn (hwæs siææl Gwdh nadhe) hadhe giffwet them ath fornempde Hollendrene skwlde sthandhe i two goordhe pa strandh, ok sagdhe sigh ath haffwa woress nodhige herres *shadfestilsse* pa for<sup>de</sup> priwilegiær som the finghe i fiordh i Køpenhampn. Tha effther sodhan breff ok *lægligheeth* som nw *affærdh* ær, ok for menigis bæsthe skwldh, særdeles ok for theris skwldh, som noordher skwle, ath the moo fonghe *wthreedningh* ok ey bliffwa altinghess *fordærwedhe* for then store thæringh the haffwe haffth i thette aar...

Også her skjønar vi vel utan vidare fleire av dei "nyare" orda, som *forfangh* 'skade', *shadfestilsse* 'stadfesting' ('bekreftelse'), *fordærwedhe* 'øydelaagd', men truleg ikkje alle, t.d. *lægligheeth* 'høve, situasjon, omstende' (jf. likevel NRO under *leilighet*), *wthreedningh* 'utrustning' og *affærdh* 'overrakt' (frå mnt. *afveren* 'absenden', sjå Kalkar under *Affærde*, bd. 1 s.15a).

Poenga her er for det første *lånorda* (frå mnt.), som ikkje alle er ordboksførte, for det andre *språkforma* generelt.

**3.2.2** Ved den første og især den siste av desse tekstane kan ein med full rett stille spørsmåla: Er dei skrivne på norsk? Og kva meiner ein med "norsk" i denne samanhengen? Tanken er at ordboksprosjektet skal hjelpe oss til svar på slike spørsmål, som kanskje til sjuande og sist må avgjerast ut frå eit skjønn. Men dette må baserast på eit breitt tekstunderlag. Då er det viktig at ein til å begynne med legg ein relativt vid definisjon til grunn. For å unngå sirkelslutningar, inkonsekvensar i kjeldeutvalet, og for å kunne avgjere (eller i det minste ta eit standpunkt til) tvilsspørsmål (t.d. ved tilfelle av språkblanding) har vi i første omgang valt delvis å gå ut frå også ikkje-språklege kriteria, på den måten at vi tek med dokument som er *skrivne* (med latinsk alfabet) *på eit nordisk språk i det som då var Noreg* innanfor perioden. Det finst saktens norsk språk også skrive utanfor Noreg, og ikkje alle nordmenn skreiv norsk (meir eller mindre "reint"). Men fordi ein i utgangspunktet ikkje kan vite for visst kva som verkeleg var, eller vart oppfatta som norsk i den tid, må ein basere utvalet på eit mest mogleg fordomsfritt grunnlag. Det gjeld å ikkje misse informasjon. Derfor er det lettare å sile ut uaktuelt materiale etterpå enn å kutte det ut, kanskje under tvil, i første omgang.

Som for alle eldre språksteg er det skriftmålet vi først og fremst, og beinveges, kan registrere. Det seier seg sjølv.

**3.3** Kjeldene er i hovudsak av tre slag: 1) diplom, 2) lovtekstar og 3) jordebøker (eigedomsg og landskyldregister). Leksikalsk sett (med tanke på vokabularet) er *lovene* på sett og vis dei enklaste, fordi mesteparten av ordforrådet er kjend frå klassisk gammalnorsk (og ordboksførd i glossaret til NgL og Fritznors ordbok (med tilleggsbandet)). Her blir oppgåva især å registrere viktigare variantar. - *Jordebøkene* er til dels store tekstar, og fleire av dei er utgjevne særskilt, såleis Aslak Bolts, biskop Øystein Aslakssons, Olav Engelbriktssons og Bergens kalvskinn. Men ordforrådet er nokså standardisert, og vil rimelegvis langt på veg bli dekt av den tredje gruppa, *diploma* (mellomalderbrev), som er langt den største.

**3.4.1** For alle tre gruppene samanlagt blir det nødvendig å gjere eit utval innanfor eit tekstkorpus på nærmare 19 000 sider. Avgrensinga er på den eine sida kronologisk og geografisk. På den andre sida er det ikkje meininga å lage nokon thesaurus over ordforrådet. Planen går ut på å registrere *alle leksem* i kjeldene, *alle bøyingsformer*, og *eit utval av variantar*. Kor mykje ein skal ta med av variantar, er delvis ei skjønnsak, men som ei rettesnor for vår praksis vil vi langt på veg gå ut frå to moment.

Det første går især på omsynet til brukarane, nærmare bestemt deira føresetnader. Vi går ut frå at dei har basiskunnskapar i klassisk norrøn grammatikk og lett kan orientere seg i vanlege handbøker og ordbøker over norrønt. Av dette følgjer at vi ikkje tek med slike variantar som normalt er registrerte i handbøkene.

Det andre momentet gjeld den generelle innleiinga til ordboka. I ei språkhistorisk skisse vil vi gjere greie for hovudlinjene i tilhøvet mellom grafem og fonem. I tillegg følgjer ei utgreiing om prinsippa for normalisering og redigering. På den måten håpar vi å fange opp ein del av variasjonen i skriveformene, og samstundes hjelpe brukarane til å finne lettare fram til oppslagsformene.

**3.4.2** Normaliseringsreglane kan ikkje fastleggjast endeleg før eit større og meir representativt materiale er gjennomgått. Men ut frå allmenn kunnskap om norsk språkhistorie har vi førebels

kome fram til dette: I alfabetet held vi **ð** og **Þ** utanfor. For **ð** skriv vi **d**, for **Þ** **th** (i trykklette pronomen og adverb) og **t**. Vokallengd er ikkje markert, med unntak av at vi har med **å**, som vi plasserer mellom **a** og **b**. Som svarabhaktivokal har vi valt **e** (mader). Dei trykklette endingsvokalane blir normaliserte **a**, **e** og **o**; dvs. at vi ikkje har vokalharmoni, endå dette er ein god del i bruk, især i første del av perioden. Sjå prøvene av redigert ordbokstekst nedanfor.

**3.4.3** Når det gjeld omfanget elles, tek vi ikkje med namn (*propria*), sidan dei aller fleste alt er registrerte særskilt i tidlegare verk (personnamn i E.H. Linds bøker, stadnamn av O. Rygh ofl. i *Norske Gaardnavne*). Omfanget av kjeldetilvisingar er eit vanskeleg spørsmål, fordi ein kan tenkje seg fleire løysingar. Eit minstekrav er at ein viser til kjelde(r) ved lemma som ikkje er ordboksført før, og elles til ord og former som av ein eller annan grunn blir vurderte som viktige eller interessante, språkleg og/eller sakleg, ut frå kriteria som må spesifiserast nærmare i innleiinga.

Siktemålet er å lage ei tospråkleg, passiv ordbok. I dette, som for så vidt er sjølv sagt når det gjeld ei såpass gammal språkform, ligg det at fraseologi og kontekst, synonym og ekvivalentar blir haldne nede på eit minimum. (Dei to siste elementa kan ein heller ikkje påstå noko sikkert om pga. avstanden i tid og språkform.) Av praktiske grunnar (behandling av lånord, plassomsyn) er norsk bokmål valt som redigeringspraak.

**3.4.4** Følgjande redigeringsprøver gjev eit inntrykk av strukturen i ordboksartiklane.

OPPSLAGSFORM	<b>vazskål</b>
ORDKLASSE	f.
TEKSTFORM	<i>waszkal</i>
TEKSTGRAMM.	nom. pl.
KONTEKST	
BETYDNING	vass-skål, vass-fat (-krukke?)
KILDE	I 321, [Voss] 1350
SMS. OG AVL.	Ja
ANDRE OPPLYSNINGER	Jf. S. Grieg: Skiftet efter Eirik Bukk på Finnen. Oslo 1934, s. 17-18.

OPPSLAGSFORM	<b>thann</b>
ORDKLASSE	pron.
TEKSTFORM	<i>they; they; Theira; them</i>
TEKSTGRAMM.	nom. pl. m.; nom. pl. f.; gen. pl.; dat. pl.
KONTEKST	
BETYDNING	den
KILDE	XXI 941, [Spangereid] 1550
SMS. OG AVL.	
ANDRE OPPLYSNINGER	

OPPSLAGSFORM	<b>heita</b>
ORDKLASSE	v.
TEKSTFORM	<i>Heither, Heitther, heitther</i>
TEKSTGRAMM.	3. pl. (?) pres. ind.
KONTEKST	
BETYDNING	hete, kalles
KILDE	XXI 941, [Spangereid] 1550
SMS. OG AVL.	
ANDRE OPPLYSNINGER	

OPPSLAGSFORM	<b>vissa</b>
ORDKLASSE	f.
TEKSTFORM	<i>vissen</i>
TEKSTGRAMM.	best. (?) akk. sg.
KONTEKST	
BETYDNING	sikkerhet, bekreftelse
KILDE	XXI 941, [Spangereid] 1550
SMS. OG AVL.	
ANDRE OPPLYSNINGER	

Av skjemaet går det fram kva slag opplysningar vi primært er ute etter. Somme (dei tre første + "Betydning") gjeld praktisk talt alle ord, andre berre visse typar. Eit eksempel på det siste er kategorien "Sms. og avl." (dvs. samansetning og/eller avleiing), der vi har tenkt å føre opp samansetningsledd og avleiingar for dei orda det er aktuelt ved, slik at vi kan få registrert samansetningar og avleiingar med visse nærmare definerte affiks særskilt (og datamaskinelt). Under "Andre opplysningar" tek vi bl.a. med tilvisingar til litteratur der vi meiner det kan ha interesse, t.d. ved oppslagsordet *vazskål* i DN I 321, der vi vil peike på eit alternativ til Fritznerns forklaring. Ein kategori som byr på særskilde problem, er den som er kalla "Tekstgrammatikk", fordi vi der må gjere ein grammatisk analyse som vi enno har for dårleg grunnlag for. Eit eksempel gjev ordet (leksemet) *thann* i DN XXI 941 (Vest-Agder 1550):

... Høy(r)dum ord oc saagom Hande band Theira som s[aa] *Heither*, aff eino Halue Steinar anundson aff andhre ... osolff torolff son Jtem Vaare *they* thaa Saa Sam satthe och venligen och vel for licktthe atth ... Jtem Lutade *they* thaa iijj syster som *heitther*... Thil ytther Meir Sanningen och *vissen* Henge vy ...

Her refererer *they* både til hankjønn og hokjønn (motsetnaden *their* : *thær* er altså oppheva). Korleis skal ein då karakterisere forma *they*? Kanskje ein ikkje bør gå lenger enn til å skrive "nom. pl." - Eit tilsvarande synkretismeproblem har vi i verba, t.d. under *heita* i same brevet. Her ser det ut til at den gamle eintalsforma er generalisert. *heiter* har sine subjekt i 3. pers. pl., men om ein derfor kan seie at forma er "3. pl.", er ikkje utan vidare opplagt. "Pres. ind." er rett nok, kanskje også "3. pers.", men "pl." blir kanskje misvisande. - Eit anna problem er forma *vissa*, også i det same brevet. Sidan det står i det formelliknande uttrykket *Thil...* *Sanningen och vissen*, der det første substantivet (*Sanningen*) står i bestemt form, har vi tolka *vissen* som bestemt form av *vissa* (eller *visse*). Men frå denne tida kjenner vi også eit substantiv *vissan* (jf. NgL V) med same tyding, og kan hende er det dette som er meint her.

I desse og mange andre tilfelle krevst det eit større materiale til underlag for analysen. Det grunnleggjande, gjennomgåande problemet er at materialet grammatisk sett er vurdert "bakfrå", dvs. frå klassisk gammalnorsk grammatikk. Derfor opererer vi med ein del distinksjonar som etter kvart vil vise seg å vere uaktuelle.

3.5 Som ein ser av redigeringsprøvene, er det lang veg fram til det formatet ei trykt ordbok bør ha. Innskrivingsformatet er førebels ein skjemadatabase, der programverktøyet er *FileMaker II* (cpr. Claris Corporation, Mountain View, Calif., USA 1988). Kvar fullstendig ordartikkel fyller ein post ("record"), som er inndelt i felt ("fields") definerte av oss. Det er ikkje ideelt, men er først og fremst valt av praktiske grunnar og for å komme raskt i gang med arbeidet. Hovudproblemet er at det er lite fleksibelt når ein må utvide eit felt, f.eks. pga. ein meir omfattande definisjon. Derfor kjem vi seinare til å vurdere andre løysingar, t.d. *HyperCard*, som er eit meir fleksibelt databaseverktøy. I alle fall reknar vi med å måtte konvertere materialet til andre format, med undertrykking av overskriftene (kategorinemningane) og omredigering av data.

4 Framdriftsplanen er ikkje heilt klar. Den vil bli avhengig av kor mange personar (ordboksredaktørar og assistentar) og kor mykje av ressursar elles som kan setjast inn i prosjektet. Gammalnorsk ordboksverk (ved Avdeling for leksikografi under Institutt for nordistikk og litteraturvitskap) har ein stab på 3 vitskapelege tilsette. Pga. andre oppgåver kan vel berre 2 av desse drive jamt med prosjektet, og då ikkje dei same til kvar tid. Dette må i hovudsaka bli eit institutt/avdelings-prosjekt, som nødvendigvis vil måtte ta ein del år. Vi tek sikte på ei litt stor (800-1000 s.) eittbands ordbok. Med tanke på fleirbruk av ordboks-materialet bør ordboka vere tilgjengeleg i to versjonar, ein maskinlagra versjon og ei trykt ordbok. Sannsynlegvis er det på den måten verket best kan gjere nytte for seg.

For å summere opp: Målet er å fylle eit hol i norsk leksikografi og samstundes byggje bru i beskrivinga av norsk mellom gammalnorsk og nyare norsk. Samtida får vurdere planen, ettertida må vurdere resultatet.

## Notar

1. Johnsen 1912.
2. Det har vore nemnt tidlegare, t.d. av Anne Holtmark: "Vi mangler f.eks. en mellomnorsk ordbok; mellomnorskens grammatikk er ikke skrevet." (1951:367). Også for islandsk er seinmellomalderen ufullstendig skildra. "... hele Perioden fra ca. 1350 og lige op til vore Dage frembyder mange interessante Problemer, som endnu ikke er blevet løst. Der findes f.E. endnu ikke nogen sammanhængende islandsk Sproghistorie fra denne Periode, men kun enkelte spredte Undersøgelser," skreiv Alexander Jóhannesson (1953:10).
3. Uthevingane i denne og dei følgjande tekstprøvene er mine.

## Litteratur

Alexander Jóhannesson. 1953. Om det islandske Sprog. *Scripta Islandica* 4 : 5-14

- Brinchmann, Christopher & Johan Agerholt (utg.). 1926. *Olav Engelbrektsøns jordebog*. Oslo
- Cleasby, Richard ofl. (red.). 1957. *An Icelandic-English Dictionary*. 2. utg. Oxford
- DN: *Diplomatarium Norvegicum*. Christiania/Kristiania/Bergen/Oslo 1849-. (Tilv. til bd. og nr.)
- Fritzner, Johan. 1886, 1891, 1896, 1972. *Ordbog over Det gamle norske Sprog*. 1, 2, 3. *Omarbeidet, forøget og forbedret Udgave*. 4. Rettelser og tillegg ved Finn Hødnebo. Kristiania/Oslo-Bergen-Tromsø
- Grøtvedt, Per Nyquist. 1969, 1970, 1974. *Skrift og tale i mellomnorske diplomer fra Folden-området 1350-1450*. 1, 2, 3. Skrifter frå Norsk Målførearkiv ved Olav T. Beito. XXI, XXII, XXVIII. Oslo
- Hellevik, Alf ofl. (red.). 1966-. *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. Oslo
- Holtmark, Anne. 1951. Norsk språkhistorie. [Melding av Indrebø 1951.] *Nordisk tidskrift för vetenskap, konst och industri. Utgiven av Letterstedtska föreningen* 27 : 365-369
- Huitfeldt, Henrik Jørgen (utg.). 1879. *Biskop Eysteins Jordebog (Den røde Bog)*. Christiania
- Hægstad, Marius. 1899. Upphavet til det norske folkemaal. (Første-fyrelesning av prof. M. Hægstad den 6/10 1899.) *Syn og Segn* 5 : 257-271
- Hægstad, Marius. 1902. *Maalet i dei gamle norske kongebrev*. Videnskapsselskabets Skrifter. I. Historisk-filos. Klasse. 1902. No. 1. Kristiania
- Hægstad, Marius. 1915, 1916, 1917. *Vestnorske maalføre fyre 1350*. 2 : 1, 2 : 2 : 1, 2 : 2 : 2. Videnskapsselskabets Skrifter. II. Hist.-filos. Klasse. 1914. No. 5, 1915. No. 3, 1916. No. 4. Kristiania
- Hødnebo, Finn (utg.). 1966. *Uvalg av norske diplomer 1350-1550*. Nordisk filologi. Tekster og lærebøker til universitetsbruk. Serie A. Tekster. 13. Oslo/København/Stockholm
- Hødnebo, Finn (utg.). 1989. *Bergens kalvskinn*. Corpus codicum Norvegicorum medii aevi. ... Quarto series vol. VIII. Oslo
- Indrebø, Gustav. 1951. *Norsk målsoga*. Utg. av Per Hovda og Per Thorson. Bergen
- Johnsen, Oscar Albert. 1912. Ord- og sageregister. I: *Norges gamle Love. Anden række. 1388-1604*. ... Første bind. 1388-1447. b. *Registre*. Christiania
- Kalkar, Otto. 1881-1885, 1886-1892, 1892-1901, 1902-1907, 1908-1918, 1976. *Ordbog til det ældre danske Sprog (1300-1700)*. 1, 2, 3, 4, 5, 6. København
- Lasch, Agathe ofl. (red.). 1928-. *Mittelniederdeutsches Handwörterbuch*. Hamburg/Neumünster
- Lind, Erik Henrik. 1905-1915. *Norsk-isländska dopnamn ock fingerade namn från medeltiden*. Uppsala/Leipzig
- Lind, Erik Henrik. 1931. *Norsk-isländska dopnamn ock fingerade namn från medeltiden. Supplementband*. Oslo/Uppsala/København
- Lübben, August. 1888. *Mittelniederdeutsches Handwörterbuch*. Norden/Leipzig. Oppt. Darmstadt 1965
- Munch, Peter Andreas (utg.). 1852. *Aslak Bolts Jordebog*. Christiania
- NgL: *Norges gamle Love*. Christiania/Oslo 1846-. (Tilv. til bd. og nr.)
- NRO: *Norsk riksmålsordbok*. Utgitt av Riksmålsvernet. 1937, 1937, 1947, 1957. 1 : 1, 1 : 2, 2 : 1, 2 : 2. Oslo
- Pettersen, Egil. 1975. *Språkbrytning i Vest-Norge 1450-1550. Språket i vestnorske skrifter ved overgangen fra mellomalder til nyere tid*. Bergen-Oslo-Tromsø
- Rindal, Magnus. 1987. *Brev frå Hamar 1350-1525. Gammalnorsk eller mellomnorsk språkform?* [Upublisert avhandling.] Bergen

- Ross, Hans. 1895, 1895-1913. *Norsk Ordbog*. Tillegg 1-6. Christiania. Opptr. Oslo-Bergen--Tromsø 1971
- Rygh, Oluf ofl. (utg.). 1897-1936. *Norske Gaardnavne*. Kristiania/Oslo
- Seip, Didrik Arup. 1934. *Studier i norsk språkhistorie*. Oslo
- Seip, Didrik Arup. 1955. *Norsk språkhistorie til omkring 1370*. 2. utg. Oslo
- Skautrup, Peter. 1944, 1947, 1953, 1968, 1970. *Det danske sprogs historie*. 1, 2, 3, 4, registre. København
- Söderwall, Knut Fredrik. 1884-1918, 1891-1900, 1900-1918, 1953, 1953-1973. *Ordbok öfver svenska medeltids-språket*. 1, 2 : 1, 2 : 2, *Supplement*. Lund
- Wessén, Elias. 1968, 1965a, 1965b. *Svensk språkhistoria*. 1. *Ljudlära och ordböjningslära*. 8. utg. 2. *Ordbildningslära*. 4. utg. 3. *Grundlinjer till en historisk syntax*. 2. utg. Stockholm
- Aasen, Ivar. 1873. *Norsk Ordbog med dansk forklaring*. 2. utg. Christiania

## Jysk Ordbog - en præsentation

På de følgende sider gives en kort orientering om *Jysk Ordbogs* redaktionsprincipper og kilder. Det vises, hvordan redaktionen i dag tager langt større hensyn til den almindelige brugers behov og forudsætninger, end man gjorde i den standsede udgivelse fra 1970'erne. Der gives til slut eksempler på ordbogens vigtigste kildetyper, så vel som på brugen af elektroniske hjælpemidler i redaktionsarbejdet.

Institut for Jysk Sprog- og Kulturforskning er oprettet 1932 og har som sin hovedopgave udarbejdelsen af *Jysk Ordbog*. Ordbogen har modtaget målrettede fondsbevillinger fra 1953.

### Medarbejdere

3 redaktører, 1 sekretær. Timelønnet medhjælp i vekslende omfang, afhængigt af fondsbevillinger etc. For tiden ca. 400 frivillige bidragydere, der besvarer spørgelister, indsender ordlister mv. (svarende til et samlet antal meddelere på ca. 3000 siden 1951).

### Samlinger

*Ordbogskartotek*: Ca. 3 mill. kartotekskort med udskrift af:

dialektlitteratur  
kultur- og lokalhistoriske værker  
indsendte ordlister etc.  
lydskrevne feltoptegnelser ved filologer.

*Spørgelistesamling*: Ca. 1300 kapsler med besvarelser af ca. 120 spørgelister, hver med ca. 40 spørgsmål besvaret af gennemsnitligt 700 meddelere; i alt svarende til ca. 3.3 mill. enkeltoplysninger (se bilag 1).

*Manuskriptsamling*: 524 manuskripter og manuskriptgrupper, fordelt på ca. 300 kassetter.

*Båndsamling*: Ca. 1120 bånd med optagelser fra ca. 550 sogne, gennemsnitligt med 25 min. tale.

*Kortsamling*: Ca. 3000 dialektgeografiske kort.



*Bibliotek:* Ca. 10.000 bind, med særlig vægt på dialektlitteratur, lokal- og kulturhistorisk litteratur, dialektologi og ordbøger.

### Publikationer

Sprog og Kultur. I-XXVIII. 1932-1978. Artikler om dialekt, folkekultur etc.  
Ord & Sag. 1 ff. 1981 ff. Årsberetning, populærvidenskabelige artikler i tilknytning til ordbogsarbejdet.

### Redaktion

I 1970 og de følgende år udsendtes 4 hæfter af *Jysk Ordbog*. Resten af bogstav A (svarende til yderligere 2 hæfter) ligger færdigredigeret efter samme retningslinjer. I 1982 besluttede redaktionen - på egen så vel som ydre foranledning - at foretage betydelige ændringer i den hidtidige redaktionspraksis. Det er siden lykkedes at sætte redaktionshastigheden væsentligt i vejret, ikke mindst ved overgangen til elektronisk tekstbehandling, og det er redaktionens hensigt sideløbende med det videre arbejde at inkorporere et revideret bogstav A i den nu edb-lagrede *Jysk Ordbog*. Om ordbogen fremover skal publiceres i bogform eller på disketter, står p.t. hen i det uvisse.

### 1970-udgaven

Redaktionsprincipperne bag de fire udgivne hæfter af *Jysk Ordbog* var helt ned i detaljerne inspireret af den i 1954 fuldførte *Ordbog over det Danske Sprog* (ODS). Den detaljerede opregning af formelle varianter, de generøse etymologiske henvisninger, den nuancerede betydningsopdeling, de præcise og vidtforgrene ordforklaringer, den massive underbygning ved citater - alle disse velkendte og træk ved ODS vil man genfinde i de to trykte afsnit af *Jysk Ordbog* i figur 1 s. 327.

Fortsat på denne måde ville *Jysk Ordbog* være blevet større end ODS. Hvad der for så vidt er forståeligt, som *Jysk Ordbog* dækker mindst ni sprogtilstande mod ODS's ene (jf. figur 2, s. 328) - men økonomisk set en vel stor mundfuld i en tid, hvor publicering var ensbetydende med en kompliceret og kostbar tryksats. Også indholdsmæssige ulemper ved ODS-konceptet talte for en ny redaktionspraksis: ODS-formen ydede ikke *Jysk Ordbogs* samlinger fuld retfærdighed, og den tog for lidt hensyn til den almindelige brugers forudsætninger og behov.

### 1982-udgaven

*Jysk Ordbogs* ny redaktionspraksis vil fremgå af afsnittet <sup>2</sup>*bytte* - *bytting*, i figur 3 s. 329. De vigtigste ændringer fra 1970-udgaven kan sammenfattes således:

1. Udeladelse af rigsmålspåvirkninger, ad-hoc dannelser og rækkesammensætninger. Gloser, udtaler, betydninger og anvendelser, der blot efterligner rigsmålet, udskydes. Det samme

gælder øjeblikksdannelser. Sætningsmedtæget kun som selvstændige artikler, når de har særlig udtale, bøjning, betydning eller udbredelse, som man ikke uden videre kan slutte sig til ud fra komponenterne.

2. Afkortet redegørelse for stof, som de jyske dialekter har fælles med rigsmålet. Betydninger kendt fra rigsmålet (og medtaget i *Nudansk Ordbog*) vil ofte kun blive gengivet som = rm. Rigsmålsnære anvendelser (herunder faste forbindelser og talemåder) lades normalt helt uomtalte, men kan fremgå af evt. citater.

3. Forenkling af artiklernes "hoved" (vedr. udtale, bøjning, etymologi). Udtaleforskelle registreres i almindelighed kun, hvor de tør anses for fonematisk relevante. Redegørelsen for udtale- og bøjningsformers udbredelse sker - om muligt - summarisk, med henvisning af de finere detaljer til ordbogens nyudviklede atlasdel (dette sidste gælder i udstrakt grad prosodiske fænomener). Etymologi gives kun for rigsmålsfremmede ordformer, som ikke allerede er behandlet etymologisk i ODS.

4. Forenkling af ordbogens betydningsangivelser. Der gives (typisk med understregning) den kortest mulige oversættelse til rigsmål; angivelsen udbygges - om fornødent - med redaktionel forklaring (uden understregning) eller - om muligt - med uddybende citater. Hvor et ord eller en form kun optræder i en enkelt kilde eller to, vil ordbogen ofte afstå fra oversættelse og lade kilderne komme direkte til orde.

5. Funktiel indskrænkning af citatmængden. Citater anvendes ikke primært til belæg af den redaktionelle betydningsangivelse. Der bringes normalt kun citater, som a) uddyber el. erstatter en betydningsangivelse, b) eksemplificerer en (ikke indlysende) syntaktisk brug, c) giver information om den rolle, det betegnede fænomen eller den betegnede forestilling spillede i almuekulturen. Mere omfattende kildemængder (herunder illustrationer) gives der blot henvisning til.

6. Præcisering af udbredelse og kildeunderlag. For hver enkelt glose, betydning eller fast ordforbindelse informerer *Jysk Ordbog* om udbredelsesområdet, ligesom der gives et groft billede af kildematerialets tæthed og alder. Hvor der er tale om rigsmålsfremmede glosser eller betydninger, ledsages redegørelsen ofte af et spaltekort.

7. Opblødning af terminologi, forkortelser og stedsbetegnelser. Sprogvidenskabelige betegnelser undgås så vidt muligt, med mindre de kendes fra skolens sprogundervisning eller nutidige opslagsværker. Der anvendes normalt kun gængse forkortelser i velkendt form. Lokalteter angives kun med fornøden præcision (i reglen mindre præcist i artiklernes betydningsdel end i "hovedet"). De anvendte stedsbetegnelser kan slås op i ordbogens atlasdel.

Fra den omtalte atlasdel af *Jysk Ordbog* vises Kort 1.4 i figur 4 s. 330. Der henvises til dette kort straks i starten af artiklen <sup>2</sup>bytte i figur 3 s. 329.

### Publicering

Det vil tage endnu en årrække at færdigredigere *Jysk Ordbog*, og så længe forekommer en publicering af ordbogen uhensigtsmæssig. Dels er en halvfærdig ordbog til ringe nytte for den

almindelige bruger; dels vil der hele redaktionsforløbet igennem være et løbende behov for detailrettelser og tilføjelser i det tidligere redigerede.

I mellemtiden vil den professionelle bruger (typisk en kollega eller studerende) kunne anskaffe en kopi af de redigerede ordbogsdele på diskette (Macintosh-formatteret), som løbende kan opdateres. Til hjælp for andre potentielle brugere kan redaktionen tilbyde diskettekopi af en liste over ordforrådet i *Jysk Ordbogs* samlinger, udarbejdet til hjælp for redaktionsarbejdet. Jfr. bilag 2.

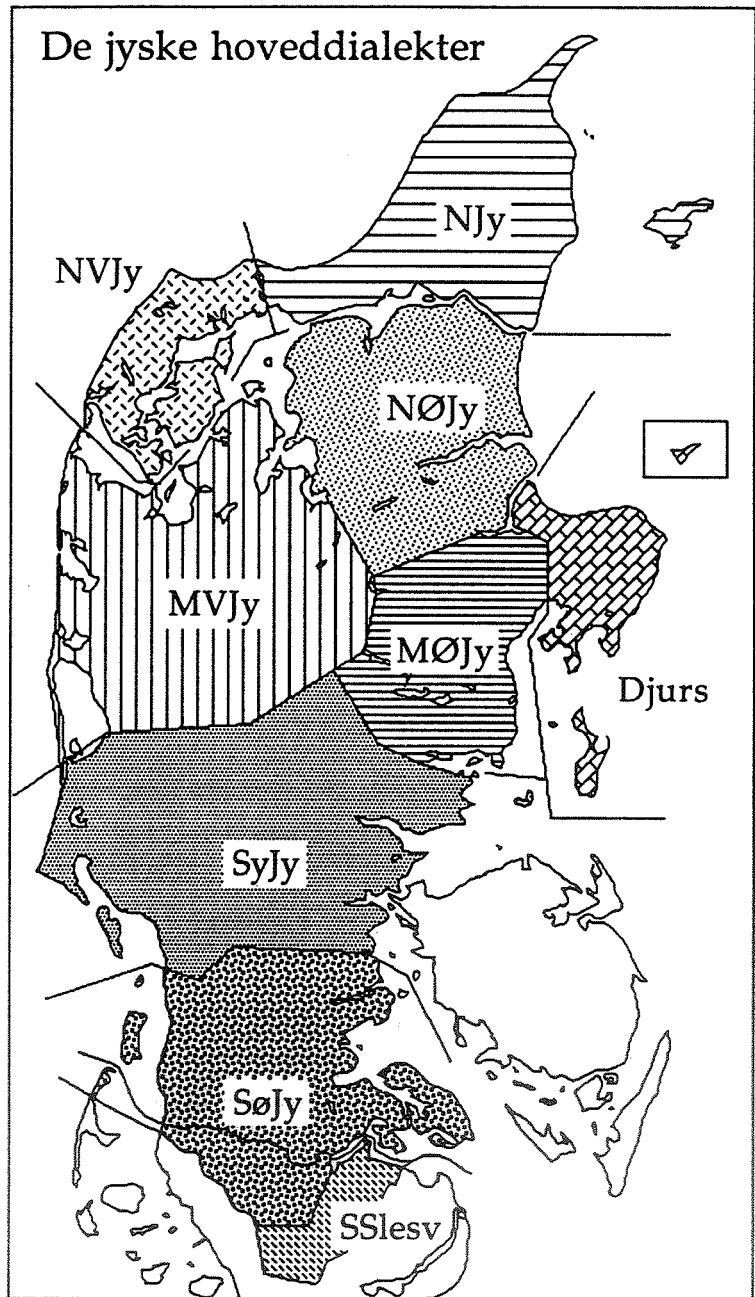
## Ordbogens kilder

*Jysk Ordbog* redigeres på grundlag af et stort, heterogent kildemateriale. Som det ses af eksemplerne i figur 5 og 6, spænder kilderne vidt: Fra håndskrevne optegnelser i starten af 1800-tallet, over trykt litteratur, til nutidige båndoptagelser. Fra knappe glossarer, over kilder hvor dialektglosser optræder i belysende kontekst, til udgivne eller uudgivne ordbøger over enkelt-dialekter. Kilder med alt fra primitiv lydgenjivelse til fuldt udviklede lydskriftsystemer.

agerhøne s *a-qərhō-n* Sejlstrup(F); *'a-qər-hō-n* HalsLæsø, AlsHindstedh; *aqərhō-ən* Voldby; *'aqər, hē-n* Tved; *'aqə(r), hō-n* Todbjerg, Krage-lund, Klakring; *'aqə, hō-n* Hundslund; *'aqər, hō-n* Hvilsager, Rær, Jelling; *'aqər'hō-n* Torsted; *'aqər, hō-n* Erslev; *'awər, hō-n* Tvis; *'awə, hō-n* Ål; *aw-ərhō-n* Darum(F); *aq-ərhō-n* Agerskov(L); *'aqə, hō-n* el *'ābā, hō-n* Vodder; *aqəihō-n* Løgumkloster; *ābāi-* el *axāi'hō-n* Haderslev; *'aqə-hō-n* Øsby; *'axə, hō-n* Åbenrå; *'axə'hō-n* Rise, Felsted; *axəhō-n* Sundeved(F), Avenstof; *'a.qəi, hō-n* Fjorde; *'axə, hō-n* Hjoldelund. — pl: *-hōns* alm; *-həns* Læsø; *-hən's* Tved; *-hōns* Klakring, Haderslev; *-hōn's* Gosmer, Endelave; *-hōns* Erslev, Fjorde. [ymda aggerhøne (Suso.139), ænvda aggerhøne (Colding.915), (pl:) aggerhøns (1555 HSmid.100), (sg:) aggerhøns (Chr.III.s bibel. 1.Sam.26.20); vn akrhœ-na; jf agerkok, agernhøne]. 1) (fugl) aggerhøne, Perdix, Briss. *a ha 'skut sɔ 'fo: 'aqə, hōns* jeg har skudt så få aggerhøns Storvorde. *'aqər, hən's di 'skratər 'sā'n mæ 'wənəm* aggerhøns de skratter sådan med vingerne Tved. *nær ə 'awər, hō-n klawər sə ɔ sejər tæ'ər — tæ'ər, sɔ væl ə'ər fry:s* når aggerhønen klager sig og siger tæer tæer, så vil det fryse Skautrup.H.1.130. *'næ'ə a 'awə, hōns di 'skri:qə sɔ fo vi 'wənə* når aggerhønsene de skrigger, så får vi vinter Ål. naar Agerhønsene skræppe tæt ved Byen, giver det en haard Tid Feilb.FH.19. || talem: *æn'tən ə'ər et en ul' ə'ər en aq-ərhō-n* enten er det en ulv eller en aggerhøne. Agerskov(L): jf.F.IV.7. 2) (insekt) mariehøne, Coccinellidæ; især: den syvplettede mariehøne. Æbeltoftegnen

8) angivende samtlige eksemplarer af en både kvalitativt og kvantitativt ubestemt (el kun vagt bestemt) klasse af personer (inden for den talendes aktuelle interessefelt): substantivisk alle; alle og enhver; alle folk; m overgange til bet 3 og 4 (bl a via ellipt substantivisering af konstruktionerne der, spec alle folk el mennesker); jf også alle-sammen, enhver; smi bet 15. mæn lival skulj ölj bādə On ā Gaməl hən ā si te men alligevel skulle alle — både unge og gamle — hen og kigge på. Tise (jf Gronb.Opt.7). *han fjætər sə 'en' hos āj*: indynder sig Tved. *'sā'n, əi'ə no 'en' ə'tə 'oləs, mən* sådan er det nu ikke efter alles Mening Nordb.Sams. *de ə 'ə'* *'jə də bæ də 'j'ə 'li: 'gət* det er ikke alle der bærer deres alder lige godt Give. i Dav (dvs på afstemningsdagen) vaar aal lig', enten det vaar Gaardmænd eller Tjennestkal' Agerskov(SjvMSkr.VI.157). — udvidet m deiktisk adv. *'āj: hə'* alle her, tilstedeværende. Ladelund. || i faste forb. alle hverén alle og enhver. Tved. *di 'kam' 'aj: 'hwə' 'jən'* de kom ... allesammen Torsted. Skvum.Mors.II.162. Rødding Røddingh(Krist.Jv.F.IX.203). Hammerumh(F). Nordby Fanø. — alle (og) enhver (vist < rm). Hundslund, Gosmer. Hadsh. — alle til hobe Gosmer. Holmsland Nvsogn. || i ordspr og talem. *dæ'ā', də jər Ol: te 'mO', jə se si:l te 'sgO* den, der gør alle tilpas, gør sig selv til skade. JMjens. Vend.249. *'ol ə 'ij 'li: 'gus* ikke alle er lige gode. PAnd.Hadsh.82. *'drek ā 'fal, 'də 'ka 'ol, mæn 'drek ā 'sto', 'də 'ka kons 'fo'* drikke og (så) falde, det kan alle, men drikke og (dog) stå, det kan kun få. Gosmer. *'tyu' 'truw'ə 'j: 'stjæ'l* tyv tror, hver mand stjæler. Give.

Figur 1. To udsnit fra hæfte 3 af *Jysk Ordbog* (udgivet 1974). Til venstre artiklen *agerhøne* s.; til højre betydning 8 af *al* pron. (artiklen fylder i alt 20 spalter + 2 helsideskort). Udsnittene giver klart indtryk af den detailrigdom, som fulgte af ordbogens ældre redaktionspraksis.



Figur 2. *Jysk Ordbog* tilstræber at dække ordforrådet inden for disse områder i perioden ca. 1700—1920.

<sup>2</sup>**bytte** v. – *byt / by.t / by\*<sup>t</sup> / ʔby(·)t* (K 1.4) alm.; *bit* °NSamsø, °Tved. – præs.: –er (K 6.2). – præt.: u.end. alm. (i Ommers og MØJy-N dog også *byttet*); *bytt* Læsø, NØ-Djurs; *by\*tt* Him-V; *bytt* Rømø. – ptc.: *byt* alm. (i Him-V dog også *by\*tt*, i Ommers og MØJy-N også *bytt*); *bit* °NSamsø, °Tved; *bytt* Læsø, NØDjurs. – passiv med reflexiv bet. (jf. bet.5): *byttes*. *byttas* / *by\*ttas* (K 1.4) Nordjy, MVJy og spredt i tilgrænsende omr. – præs.: u.end. – præt. og ptc.: –t.

1) = **rm.** || bytte noget –hen. – (navne på lege:) bytte –'gårde, bytte –'huse, bytte –'koner. || (spøg.:) bytte 'løgn = spørge nyt. °°Haverslev. – (spec.): bytte et ord = skændes. *A haajj haalt ow Else, de ku a no faanim, faa wal haaj vi ... byt et Uwe en Gaang imell, men de kam vi godt uj ow =* jeg havde holdt af Else, det kunne jeg nu fornemme, for vel havde vi skændtes en gang imellem, men det kom vi godt ud af. Thomaskjær.PK. – bytte 'gær. *når man skulde til at brygge, kunde man sige til Pigen: 'no ska do 'uj' å bit 'gja'r 'pi'q (nu skal du ud at bytte Gær, Pige); ... når man var færdig med at brygge, stod Gæren i 'gja'r påtøn til den næste Gård skulde brygge.* °°Nsamso.

2) = **dele** [Vends-N, NVJy, Sall, Hards-N, Fjends-N; spor. i øvr. Nord- og Midtjy; jf. *skifte*] *feskørøn by.t feskil, som di ha få.ñt i 4 1/2 lø = fiskerne delte fisken, som de havde*



<sup>2</sup>bytte 2 = dele

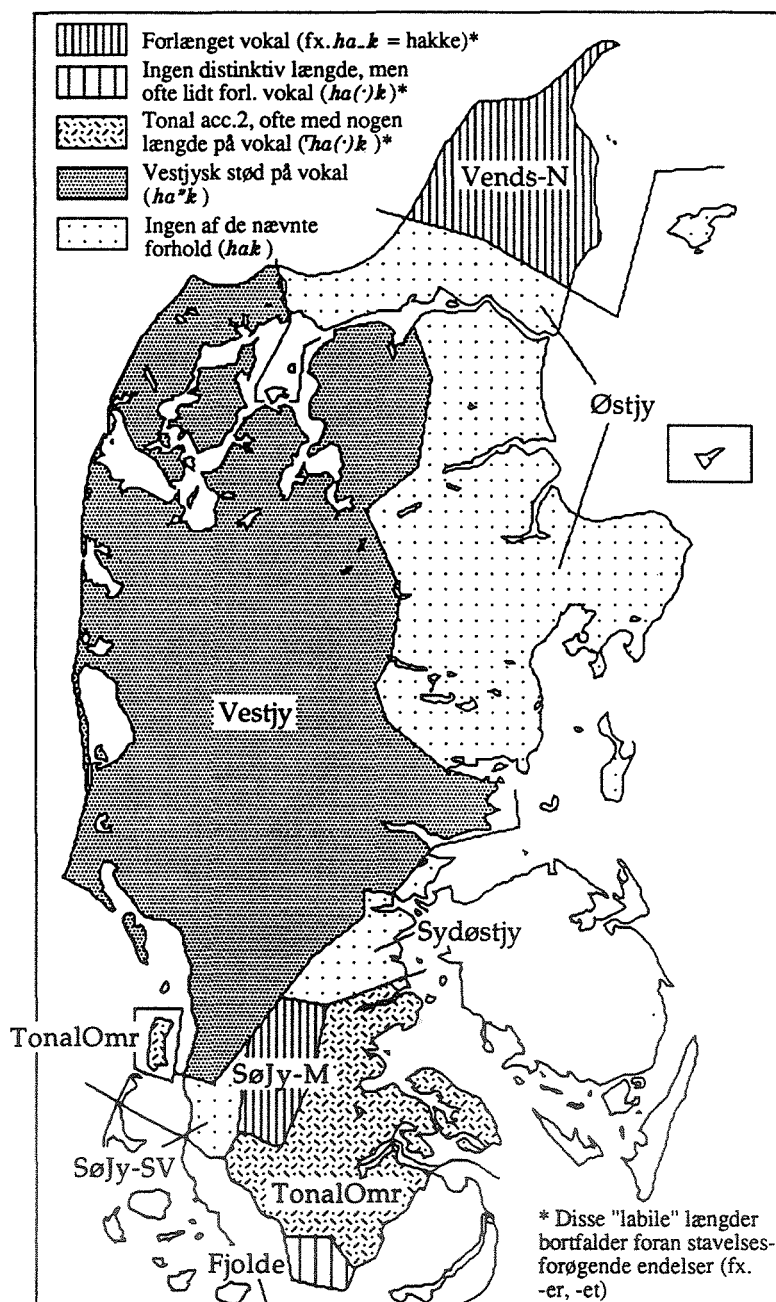
*fanget, i 4 1/2 -lod (ved pilkefiskeri fra Løkken).* AEsp.VO. 'hæ'r 'ær' øn 'då'lor i ka by\*<sup>t</sup> = her er en daler, I kan dele. °Torsted. *ska wi by\*<sup>t</sup> øn 'æ'ðal =* skal vi dele et æble. °Torsted. || (i særl. anv.): = **udskifte tidligere fællesareal**. *de wa fa'r markøn blöw byt =* det var før marken blev

delt (ved udskiftningen). °°Vorning. *Mange år efter at marken var byttet op, kunde man høre mændene fra Hvorslev gå ude på marken og rasle med landmålerkjæden* (nemlig som gengangere). °HoulbjergH (Krist.JyF.VIII.-199). – **bytte ad** = d.s. [Hards-N] *vi byt æ kaag aj =* vi delte kagen. SkodborGH. *da de har faaet den* (dvs. en pot brændevin) *byttet ad, er Manden jo blevet lidt rusende.* VandfuldH (Krist.BK.93). – **bytte om** = dele om, fordele. *bytt' Ma om = dele Mad rundt.* °Mors (Schade.41). || (iflg. F. også) = **fordele, uddele** [Østjy] *to karle vælges te å byt gajlam = uddele -gadelam.* Testrup (F., NingH?) – hertil (i dagbogsopt. 1767): *Den 28. blev hovmødding byttet* (vel = fordelt, kørt ud). °Als (ChrHans.Elstrup.21).

**byttel** s. – (I/IK 4.8:) *bytt* – fk. [Sønderjy-SV; i sms. *mel·byttel* også spor. i øvr. SøJy samt i GørdingH; jf. *sakkuk, sætte'kage*; < nty. *mehlbüdel* el. ty. *mehlbeutel* = budding, opr. kogt i *beutel* = lærredspose; jf. *Dialektstudier.V.261*] = **mel·byttel**. || hertil: **byttel·dåse** s. = buddingform. EKWestergaard.DE.14.

**bytte·vogn** s. = **vogn uden forspand**; den læsses, medens spandet kører med en tidligere læsset vogn. °ØHanH (AEsp.VO.).

**bytting** s. – 1) = **rm.** (om barn af underjordiske) [spor. i Nordjy] 2) = **lille vanskabt person**, som er lidt til en side. °AEsp.VO.



Figur 4. Kort K 1.4 fra *Jysk Ordbogs atlasdel*: Længde-, tone- og stødforhold i accent-2-ord med kort vokal + p, t, k.

Om, N. S. Betænkelighed, Vankelmød: der er en Om ved; der maae ingen Om være ved  
 Løben, Løftet, Købet etc.!

Om ved; der maae ingen Om være ved / Af,  
 Løben, Løftet, Købet etc.!

Om ved; der maae ingen Om være ved / Af,  
 Løben, Løftet, Købet etc.!

Om ved; der maae ingen Om være ved / Af,  
 Løben, Løftet, Købet etc.!

Om ved; der maae ingen Om være ved / Af,  
 Løben, Løftet, Købet etc.!

Figur 5. Indberetning fra by- og herredsfoged Laurids Fogtman ca. 1810 om "provindsielle Ord" på Vardeegnen. Overstregningerne skyldes Chr. Molbech, som udnyttede adskillige indberetninger af denn art i Dansk Dialect-Lexikon (1841).

Rårup i Bjerrø herred

et bæn, o bæn (flæstet) en bæn,  
 bæn (Køjsnote)

flæd, læ gla  
 skæd, skæd - skæ  
 hiti = hie - hie = heid

I ku var en søregen lyd  
 hvor se synes at mødes  
 et stæft frales (2)

smel smel smit's Pütz  
 daw daw } de mjolk } ren, de  
 } de hie } di gøt-gøt }  
 } di hie } di fløt fløt }  
 } } r Tungelyd

fæd fæd bla bla  
 lam lam fol fol  
 høst høst

Figur 6. De første filologiske optegnelser af jysk blev foretaget af K.J. Lyngby - med henblik på en aldrig færdigskrevet jysk grammatik. Prøven her (fra 1856) stammer fra Rårup ved Vejle. Den viser både Lyngbys systematiske spørgeteknik (lydkorrespondenser, bøjnings-skemaer) og hans spontane iagttagelsessevne (dialektens eu i ord som ko).

Aflæsning:  
 Om, N. S.\* Betænkelighed, Vankelmød: der er en Om ved; der maae ingen Om være ved  
 (Aftalen, Løftet, Købet etc.)

Omsæt, at gjøre, dvs. at kunne godt indrette sine Ting; f.E. en Kone veed at gjøre Omsæt, naar Hun i en Hast faaer uventede Fremmede.

Oplod, N.S. det Øverste Halve af en Serk.

Opskiør ell. Opskiær, N.S. Høst Gilde.

\* nomen substantivum, dvs. substantiv.

## 42

„X vil ha Dæ, aa plej Dæ aa heh Dæ po din gammel Daw, de haa Do faastylt.“ — „Hohoho!“ søh han; de hør- ed han. X bedrow ham heller ett: X had ham po syw- end Der, aa han had bejr Daw end X haad sje, saa han bestelt slet ett Ant, end X rej po ham engaang imell te Rjovstajen, aa i Sæhti aa i Høsten ku han gver et Beed ivessonda. — Men vi ka ett loww te ewig Zi — Jens maat aase afej. — X haad engaang wo hjemmenfra en si Daw, aa da X kam i Goeren, wa der iff et lowend Minnist aa sie hwerten op elle nier, undtawn mi gammel Mowr, da stod aa ragt i Grødden po Slaastien. „Gus Frej!“ sø X, „hudden aret fat? hvor ar aall di Anner?“ „Di ar uhd i Tøpsten ve Messingjens“ sø hun, „faa han ve te aa stæmt.“ X gif derud. Di stued om ham baade Pelle, aa behgi wo Bønh, aa Kaalen aa Pigen, aa Jens toe imell dem po hans Eih: „De a not stit mæ Dæ“ sø X. Han letted hans Pomb, aa soe saa naalle te mæ, aa sokked, aa loe Pordet nier igjen, strækked hans Bien aa dsh. — „Pelle!“ sø X, „han stul indda be- graves let hebdede; X vil ikke træl Rjomlen aa ham faar aalle de; aa han sta jores i Kalgoren, faar X vil hwerten ha Støyen eller Punn te aa snahg i ham.“ — Saa kam han da aasje te aa leagg unne den stuer Avvil, som mi Faar had stammed. — Swægaang X sie Træt, tinket X po Messingjens.“)

\*) Der ere endnu (1842) flere levende af dem, som kjend- te Messingjens i levende Live, og hans Eftermæle lever, saavide jeg veed, mellem de flestvigske Kyrasjerer.

Figur 7. Banebrydende for dansk dialektlitteratur blev Steen Steensen Blichers *E Bindstouw* (1842) - med en række digte og fortællinger på jysk. Mest populær blev nok fortællingen om *Messingjens*, der slutter som vist her. Det ses, hvor tæt Jysk Ordbog har excerperet *E Bindstouw* (markeret ved understregning).



265. I egnen ved *Oksenvad*, Haderslev amt, kaldes det sidste neg æ *fisneg*, og pigen, som binder det, kaldes æ *fismoe*, hendes karl æ *fisfæ*. Den sidste kjærv på marken kaldes æ *fiskjære*, og den bliver pyntet og stillet op midt i gården, hvor ænder og gjæs har den at gjøre sig til gode med. Ved høstgildet skal æ *fisfæ* og æ *fismoe* sidde ved skorstenen (ildstedet) at spise og drikke.

Mikkel Sörensen.

266. Den, der her binder det sidste rugneg, får rugstodderen, det sidste bygneg bygstodderen, det sidste havreneg havrepiggen.

Den, der i Vardeegnen fik det sidste neg, når der sattes korn sammen, fik enkemanden.

Lærer Lauritsen, Sulsted.

Figur 8. 1867 begyndte Evald Tang Kristensen en omfattende indsamling af folkeminder - hovedsagelig fra Jylland. I figur 7 ses et par afsnit fra *Det jyske almueliv*, Tillægsbind I (1900). Mange af optegnelserne er spækket med dialektgloser, som er excerperet til *Jysk Ordbog* (her markeret ^).

Foran nævnte Skovfoged Christian Hansens Bo i Linaa (fra 1753) kan formentlig anføres som et velforsynet jordløst Husmandshjem.

Deri nævnes:

Sølv: 5 Skeer. — Kobber: 3 Bryggerkedel. — Malm: 1 Par Lysestager. 1 Morter med Støder. — Tin: 3 Fade. 1 Par Lysestager. 1 Potageske. 2 Thepotter. — Messing: 2 Kedler. 1 Fyrbækken. 1 Rivejern. 1 Lysestage. — Jernfang: 1 Kakkellovn. 4 Gryder foruden det alm. — Træfang og dertil 1 Stue Uhrverk med Fotteral. 1 Speil. 2 Perremie<sup>1)</sup>. 13 Stole. 2 Spinderokke. 1 Puster. 1 Rulle. 1 Liigkiste af Fuhr (Fyr).

<sup>1)</sup>? [Forf. ved ikke, hvad Perremie er]

Figur 9. Den historiske interesse i sidste halvdel af 1800-tallet førte til omfattende publicering af beretninger og dokumenter, som belyser de danske egnes kulturhistorie. Det viste udsnit af *Aarhus Stifts Aarbog 1916* i figur 9 vedrører et skifte fra 1700-tallet. Udgifveren beklager, at han ikke forstår glosen *Perremie*; *Jysk Ordbog* vil sikre, at en fremtidig læser henvises fra denne grove skrivemåde til: *pyramide* s. (= amagerhylde).

man 'so, da wi blew 'let 'gamlar, so  
 blew min 'mow'r 'sy'g; o 'so war ed  
 'ma, de sku 'pas de 'he:la, for a  
 hað 'gát ,nok 'hat en 'sestar, man 'hon  
 war 'dej. o 'so skul a 'et rän te 'foir,  
 man 'nä'r a gik te 'foir, so skul a 'et hä  
 'low te o 'go', ujan a sku hä med 'berytöw'  
 'ma ma, for 'hæf:əs so war ed jo 'galt.

Figur 10. Båndoptagerens fremkomst har gjort det lettere at optegne sammenhængende tale. I figur 10 ses udskrift af en båndoptagelse fra øen Anholt 1956, begge dele ved Niels Åge Nielsen. Udskriften er blevet tæt exciperet til *Jysk Ordbog*, især med henblik på gode mundrette citater. En tilsvarende udnyttelse af instituttets øvrige båndsamling er desværre økonomisk og tidsmæssigt uoverkommelig.

*Oversættelse:*

Men så, da vi blev lidt glamlere (= ældre), så blev min mor syg; og så var det mig, der skulle passe det hele, for jeg havde godt nok haft en søster, men hun var død. Og så skulle jeg ikke rende til får ... [selvrettelse:] men når jeg gik til får (dvs. vogtede får), så skulle jeg ikke have lov til at gå, uden jeg skulle have mit bindetøj (dvs. strikketøj) med mig, for ellers var det jo galt.

Figur 11. Systematisk undersøgelse af en dialekts ordforråd kræver årelang indsats af en lokalkendt kapacitet - men ikke nødvendigvis med universitetsuddannelse. Blandt *Jysk Ordbogs* sikreste kilder er de store ordsamlinger ved lærerne Torsten Balle (Thy) og Jens Skytte Andersen (syd for Skive). Ordsedlerne i figur 11 er karakteristiske for disse samlingers omhu i lydangivelse og ordforklaring:

dagskær ----- s -----  
 'daw:skjær  
 " - varmegvirkning af dagslyset (ikke direkte sollys)  
 -----  
 om vinteren, Denne varmegvirkning kan få sneen  
 på sydsiden af et tag til at smelte, selv om luf-  
 tens temperatur er noget under frysepunktet. Man  
 siger da: de 'dre 'pør fræ ø 'hū:s , 'de æ 'nåk ø  
 -----  
 'daw:skjær.

*bål'k; sb. - bulke -; flt. bål'g r.*  
*... muret sgjæl-rom' (skillerum), f.eks.*  
*som adskillelse mellem stier i gamle*  
*stalde nående ca. halvvejs til loftet.*  
*- æ sow' hå vælt æ bål'k in' te 'smo' gri*  
*s<sup>1</sup>; - overf.: vi mo ves hæl r re-s en bål'k imæl i an r tåw 'slaws' brøj r!<sup>2</sup>*  
*(få jer skilt fra hinanden). (Oversættelse: <sup>1</sup> soen har væltet balken ind til smågrisen; <sup>2</sup> vi må vist hellere reise en balk imellem I andre to slagsbrødre).*

**Aflæsning:**

*bål'k; sb. - bulke; - flt. bål'g r.* - muret *sgjæl-rom'* (skillerum), f.eks. som adskillelse mellem stier i gamle stalde nående ca. halvvejs til loftet. - *æ sow' hå vælt æ bål'k in' te 'smo' gri s<sup>1</sup>*; - overf.: *vi mo ves hæl r re-s en bål'k imæl i an r tåw 'slaws' brøj r!<sup>2</sup>* (få jer skilt fra hinanden). (Oversættelse: <sup>1</sup> soen har væltet balken ind til smågrisen; <sup>2</sup> vi må vist hellere reise en balk imellem I andre to slagsbrødre)

- /384/ [538] og (konJ) , og (konJ) de (pron-ubest) skulle (v-præt)  
 /385/ [539] Jo (adv) være (v) så (adv) stram (a-pl) , at (konJ)  
 /386/ [540] når (konJ) - ned (adv) ad (præp) en (art-m) bakke (s-m)  
 /387/ [540] sådan (adv) at (konJ) \*svingeltræ (s-pl-bf) de (pron)  
 /388/ [542] skulle (v-præt) Jo (adv) ikke (adv) tage (v) imod (adv)  
 /389/ [542] \*hase (s-pl-bf) af (præp) de (pron-oblik) . Ja (interJ)

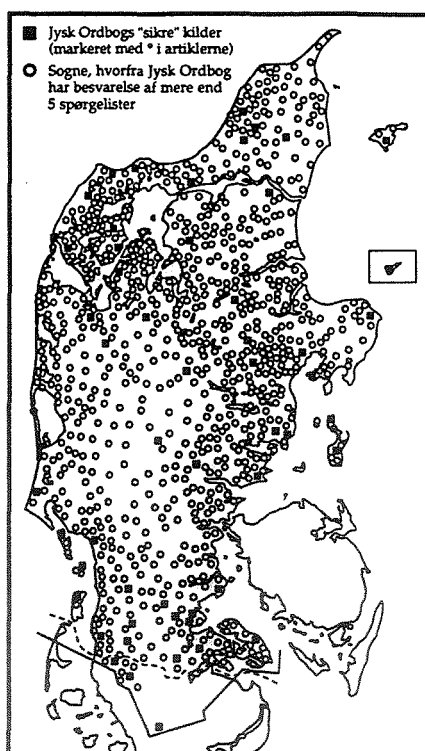
**Oversættelse:**

og de (dvs. skaglerne) skulle jo være så stramme, at når (det gik) ned ad bakke ... sådan at svingeltræerne de skulle jo ikke tage imod haserne af dem (dvs. hestene), ja.

Figur 12. Ca. 100 bånd fra instituttets samling er udskrevet i normaliseret form. Tekstens ord gengives som glose + ordklasse + bøjningsform (samt \* for mere sjældne dialektord). Normaliseringen muliggør elektronisk behandling, f.eks. alfabetisk og topografisk listning af ord og bøjningsformer.

ORD	PRÆPARERING	LINIE	TRIP	BÅND	SOGN
SVINGE					
V		665	547	460	1229A
V		155	211	334	1527
V		107	151	344	1535
V-PRÆT		852	680	63	1994
V-PRÆT		349	389	307	2501
V-PRÆT		381	414	307	2501
V-PTC		136	183	99	1182
SVINGELTRÆ					
S-PL-BF		387	540	632	1259
SVINGLE					
S-PL		75	105	371	2153

Figur 13. Udsnit af en alfabetisk ordnet liste over de \*-markerede ord i samtlige instituttets båndudskrifter. For den enkelte ordform er anført (kolonnevis i denne rækkefølge): linjenummer i udskriften, triptæller-reference til båndet, båndnummer, topografisk sognenummer. Listen anvendes som supplement til det øvrige redaktionsmateriale og viser, hvor en udtale kan efterhøres i påkommende tilfælde.



Figur 14. For hele det centrale ordforråds vedkommende hviler *Jysk Ordbogs* oplysninger først og fremmest på to hinanden supplerende kildetyper: 1) de "sikre" kilder (med dækkende lydskrift og grundige ordforklaringer, optegnet af filologer eller trænede lokalfolk), 2) det store antal spørgelistebesvarelser (som på trods af mangler og evt. fejl i den enkelte besvarelse tilsammen giver et sikkert billede af ords, udtalers og betydningers geografiske udbredelse).

## Bilag 1

29 JAN. 1990

JYSK ORDBOG  
SPØRGELISTE DB

1286 4

<p>1. Kendes i Deres dialekt ordet at <u>dominere</u>? Hvordan udtales det? (domenijr, domenjer, domenger, el. anderledes) Hvad betyder ordet? (larne, skælde ud, bande, el. andet) Giv eksempler på ordets brug.</p>	<p>å dommä'ni'r ; hanj 'war let dommä-'ni'rän^ = han var lidt dominerende; men: hanj 'banj.t å tu'ni'rä = han bandte og turnerede ?/tórdnede ? (skældte ud)</p>
<p>2. Hvad hedder i Deres dialekt udtrykket "jo mere <u>des</u> bedre"? (jo mejer des bejer; es mijer es beer; je mee je bee; i mier i bæjer; el. anderledes) Skriv udtalen af hele udtrykket.</p>	<p>jo 'mi.r jo 'bæj.r ; des 'mi.r des 'bæj.r ; jo 'källär (å som i: op) jo 'bæj.r = jo værre jobedre ; jo 'wæ.r des 'bæj.r = jo værre des bedre ; (og vist også andre kombinationer efter humør og behov)</p>
<p>3. Hvad hedder <u>seletej</u> i Deres dialekt? (dræt, dræht, dreet, tæw, tæw, ty, el. andet) Gengiv ordets udtale i sætningen "Se nu at få lagt <u>dræt</u> / <u>tøj</u> på hesten!"</p>	<p>'sij' 'no å fo lå 'drät po 'häjstän</p>
<p>4. Kan <u>dræt</u> i Deres dialekt bruges i én eller flere af følgende betydninger? (a) foer (i tøj)? (b) spor efter noget (f.eks. efter en harve)? (c) indhaling af vod? (d) arbejdsperiode (før der holdes pause)? (e) opdræt (af kvæg)? (f) nakkedrag? Giv eksempler på brugen.</p>	<p>a). 'drät - i tøj b). kendes ikke c). kendes ikke e). opdræt = 'åpfenneng (opfødnung ; eks.: wi føj. 'fäm 'kal. 'åp åm 'o'rä = vi opdrættede fem kalve om året d). en arbejdsperiode = ä 'bi' f). et nakkedrag huskes af 90-årig som: ä 'drag</p>
<p>5. Kendes i Deres dialekt ordet at <u>djævle</u> el. at <u>djævles</u> (i en forbindelse som "bande og djævle", "djævles med nogen el. noget")? Hvordan udtales ordet? (døwwl, dæwles, dywles, el. anderledes) Hvad betyder ordet? (bande, skråle op, kives med, trækkes med, tumble med, el. andet) Giv eksempler på ordets brug.</p>	<p>'no wel a 'ejk 'däw.läs 'mä'ä 'länggär = nu vil jeg ikke mase/ tumble/trækkes med det længere :</p>

Spørgelistebehandling indsendt 1990 af Anders Fisker (Hanherred). Den anvendte lydskrift dækker i næsten alle henseender dialektens fonematiske struktur.

## ORDLISTER

## 1

opslagsform	ordkl.	betydning, /brugssfære/	udbredelse	kilder	spørgel.
1.basse'ralle	s	larm, gilde	J	3	
2.basse'ralle	v	larme	NØ	1	
bass+eret	adj	befængt med lus	NN,NØ	2	
basse'rok	s	/fugl/	NN,MV	2	
basse'rolle	s	snurretop	NØ	1	
basse'rone	s	stort gilde	SS	1	ÅÅ29
basse*skaff+er	s	(ved gilde)	MØ	1	
basse*spyd	s	bjørnespyd	MV	1	
bassevak	s	/spøg/ kæleord for barn	NN	1	
1.bass+i(e)	v	1) gå med besvær	NN,NØ,MV	2	ÅÅ27,Z12
1.bass+i(e)	v	2) søle sig til, svire	MV	2	
2.bass+i(e)	v	1) gasse sig	N,M	3	
2.bass+i(e)	v	2) rode, lave uorden	MV	1	

## 2

aflædn.end.	opslagsform	ordkl.	betydning, /brugssfære/	udbredelse
+i(e)	af*græss+i(e)	v	afgræsse	MØ
+i(e)	2.bass+i(e)	v	1) gasse sig	N,M
+i(e)	2.bass+i(e)	v	2) rode, lave uorden	MV
+i(e)	1.bass+i(e)	v	1) gå med besvær	NN,NØ,MV
+i(e)	1.bass+i(e)	v	2) søle sig til, svire	MV
+i(e)	3.bass+i(e)	v	/leg/ (i remse)	MV
+i(e)	duss+i(e)	v	gå med små skridt	M

## 3

/brugssfære/	opslagsform	ordkl.	betydning	udbredelse
/fugl/	bakke*svale	s	/fugl/	N,M
/fugl/	bakke*ugle	s	/fugl/	NN,NV
/fugl/	banke*fugl	s	/fugl/	NØ
/fugl/	bartel*fugl	s	/fugl/	MØ
/fugl/	basse'rok	s	/fugl/	NN,MV
/fugl/	bassumi	s	/fugl/	NN

Udsnit af: 1. Liste over ordforrådet i Jysk Ordbog. 2. Liste over forekomsten af afledningsendelser; +i(e) og +ie optager i alt 170 linjer. 3. Liste over brugssfærer; fuglenavne (præpareret /fugl/) optager i alt 1196 linjer.

Som arbejdsredskab for redaktionen og til orientering for andre interesserede er der udarbejdet en liste over ordforrådet i *Jysk Ordbog* (pkt 1). Listen rummer kortfattede betydningsangivelser (i visse tilfælde blot en henvisning til betydnings- el. brugssfære, markert / /), desuden summarisk angivelse af udbredelsesområderne samt kildernes mængde (1 = 1 kilde, 2 = 2-10 kilder, 3 = 11-50 kilder, 4 = over 50 kilder); endelig angives det, om ordet har været efterspurgt på spørgeliste. Opslagsformerne i listen er præpareret sådan, at der har kunnet udarbejdes særlige lister over afledninger og 2. sammensætningsled såvel som over betydnings- og brugssfærer. I pkt. 2 ses et udsnit af listen over afledninger på -ie (i alt 170 linjer), i pkt. 3 et udsnit af listen over fuglenavne (i alt 1196 linjer).

Tor Erik Jenstad

## Presentasjon av prosjektet "Ordbok over trøndermåla"

Artikkelen gjer greie for bakgrunn, status og framtidsplanar for "Ordbok over trøndermåla". Dette første større regionale ordboksprosjektet i Noreg skal dekkje dialektane i Trøndelag og på Nordmøre. Eit hovudpunkt er å dele prosjektet i to, med ei folkeutgåve av ordboka på kort sikt og ei meir utførleg vitenskapleg behandling på lengre sikt.

### Bakgrunn og status

"Ordbok over trøndermåla" (OTM) er tenkt å bli ei regional ordbok over dialektane i Trøndelag og på Nordmøre. Det er eit prosjekt som etter ein bra start diverre har vore hardt utsett for underernæring, og som av den grunn har vorte nokså hemma i utviklinga. Men det står truleg ikkje verre til enn at ein snuoperasjon vil kunne få pasienten på føtene att. Denne snuoperasjonen lyt først og fremst vera av økonomisk slag, og det eg vil seie med dette, er sjølvstøtt at finansieringssituasjonen for prosjektet i lengre tid no har vore temmeleg håplaus. Men manglande midlar er jo ikkje noko ukjent fenomen for leksikografar.

Prosjektet vart starta i 1981 under leing av Reidar Djupedal og Arnold Dalen. Frå og med våren 1987 vart det etablert eit samarbeid med det som den gong var Norsk leksikografisk institutt, Universitetet i Oslo, slik at ei av dei stillingane ein der fekk ekstra løyvingar til frå og med dette året, skulle brukast til å leggje til rette for *Norsk Ordbok* (NO) tilfang innsamla på OTM-prosjektet. For meir utførleg framstilling av OTM-prosjektet sjå elles Jenstad 1984 og 1989. Arbeidsstaden og samlingane til OTM ligg ved Nordisk Institutt, Universitetet i Trondheim.

Innsatsen i prosjektet har vore konsentrert om innsamling og ekserpering. Den redigeringa som har vore gjort, må sjåast som reinte førebuaende, og det har til no ikke vorte publisert noko. Vi har samla tilfang ved å sende ut spørjelister til informantar rundt ikring i kommunane (på det meste om lag 200 personar), og dette kontaktnettet er enno nokolunde intakt. Arkivet som er bygd opp, er nokså omfattande, og vi meiner det er forsvarleg å ta til med redigering/utgjeving no, særleg dersom vi kan halde kontakten med heimelsmannsnettet parallelt med at redigeringa skrid fram, slik at vi kan konsultere dei der vi oppdagar hol i tilfanget.

Dialekteleksikografien er ei grein av allmenn leksikografi, og i prinsippet same sak. I praksis er det like vel visse spørsmål som teiknar seg klårare når ein arbeider med dialektalt ordforråd. Eit regionalt prosjekt som dette, der ein skal dekkje eit spekter av ganske varierende mål, byr i tillegg på sine heilt særigne problem. Ein del slike har reist seg for oss som har arbeidd med OTM med tanke på publisering, og det er det resten av artikkelen skal handle om.

## Ei folkeleg dialektordbok

Ein tanke som har vorte meir og meir sentral for oss etter kvart, er å dele prosjektet i to, ein folkeleg og ein vitenskapleg del. Ei folkeleg utgåve av ordboka kunne utarbeidast etter måten snøgt, medan arbeidet med ein vitenskapleg versjon ville måtte ha eit meir langsiktig perspektiv. Vi trur vi tør påstå at OTM har hatt god folkeleg kontakt. Innsamlingsarbeidet har gjeve oss vitjing av folk som elles neppe ville ha sett sine bein innafør eit universitet. Ein slik kontakt er til vinning for begge partar. Ansvarer i høve til heimelsfolka synest eg er eit sterkt argument for å skunde på ei utgjeving. Mange av dei er eldre menneske, og dei burde få sjå resultatet av innsamlingsarbeidet sitt i ein større samanheng før det er for seint. Ikkje få av dei som var med frå starten, er borte no. I tillegg har vi hatt ganske god eksponering av prosjektet i media. Samla sett gjer dette at vi føler eit forventningspress om å få ut noko. Tilfanget har svolle opp og vorte så stort at ei fullstendig vitenskapleg behandling av det vil ta svært lang tid, i alle høve med dei utsiktene det er for ressurstilførsel no. Ei slik utgåve kan ein rekne med vil ha interesse først og fremst for forskarar, og dei vil kunne få tilgang til materialet kor som er. Men ein folkeleg versjon meiner eg at vi skulle makte å få ferdig innan ei rimelig tidsramme, t.d. eit par år.

Kven tek vi sikte på å nå ved ei folkeleg trønderordbok? Når ein ordboksmakar skal mottakarorientere seg, famlar han lett i blinde, sidan vi eigenleg veit svært lite om korleis ordbøker faktisk blir brukte. Det gjeld ikkje minst dialektordbøker, og særleg behova til vanlege brukarar. Tanken vår er å vende oss til den interesserte amatøreren i von om at den typiske informanten vår også skal bli brukar. Ordboka må altså utformast slik at ho blir allment forståeleg, og leseleg for folk flest. Samstundes må det som står der, vera forsvarleg og rett; ei popularisering fører ikkje til at ein kan unndra seg vitenskapleg kritikk. Ein må også rekne med at forskarar vil bruke ei folkeutgåve, mottakargruppa vil truleg i det heile bli nokså heterogen. Men krav til fullstendig og uttømmmande framstilling, utførelg dokumentasjon osv. lyt det nok slakkast på.

Korleis skal vi så greie å gjera ordboka folkeleg? Heilt allment kan ein vel slå fast at ho ikkje må bli for stor og uhandterleg. Det gjer at ein lyt skjera ned der ein kan, og at det vanlege plassproblemet i leksikografien her blir enda meir presserande. Ved større ordbokstiltak gjennomfører ein gjerne ein slankeprosess for å komprimere til hand- og lommeutgåver. Men da går ein jo ut frå den fullstendige versjonen og skjer i den, medan vi altså tenkjer oss å gå den omvendte vegen og laga den forkorta versjonen først. Arbeids- og tenkjemåten kan såleis ikke bli heilt parallell. Eit overordna prinsipp lyt likevel bli å uttrykkje seg kort.

### Formelle sider

Forma på *oppslagsordet* er eit tradisjonelt problemområde i dialektleksikografien. Ein kan i prinsippet tenkje seg tre moglege løysingar:

- 1 å setje dialektforma som oppslagsform
- 2 å setje ei riksspråksform som oppslagsform
- 3 å setje ei eldre (heimla eller rekonstruert) form som oppslagsform

Den siste ser eg ikkje som aktuell i det heile i ei ordbok som vil vera folkeleg. Det ville bli altfor vanskeleg å finne fram for den som ikkje har spesialkunnskapar i språkhistorie.



Løysing 1 er fullt ut tenkjeleg i ordbøker over einskildmålføre. Denne kan kombinerast med ei normalisert form i parentes, slik det er gjort i Killingbergtrø 1983. Dermed får ein også i konsentrert form sagt ein god del om opphavet til ordet. For den som ikkje snakkar dialekten sjølv eller har god kjennskap til den, kan ein ha eit eige register med normalformer med tilvising til sjølve den leksikalske delen. Men for oss som skal dekkje eit dialektområde med såpass stor indre variasjon, blir prinsipp 1 vanskeleg. Det blir det også dersom ein skal velje *den mest utbreidde dialektforma* som oppslagsform, ein utveg som har vore nytta i fleire lågtyske ordbøker (Niebaum 1986:131).

Dermed står vi att med løysing 2, og det blir nok ei riksspråkleg oppslagsform vi kjem til å lande på. Den norske rettskrivinga er da også såpass romsleg at det i mange høve let seg gjera å velje former som ligg nokså nær uttalen og har greie korrespondansar til den. Vi vil streva etter å bruke oppslagsformer som høver med systemet i trøndermåla, t.d. å gjennomføre kløyvd infinitiv.

Noko av den største variasjonen i trøndermåla finn vi i jamninga, med ytre former som *vækka*, *færra* mot indre, austlege som *vukku/vokko* og *fårrå*. Her ser eg ikkje anna råd enn å føre alle variantane under same normaliserte oppslagsform (altså *vækka* og *vokko* under *veke*). Dette vil ein måtte kombinere med ein omfattande bruk av tilvisingar frå dialektform til oppslagsform, sidan ein ikkje utan vidare kan rekne med at brukaren skjønner samanhengen mellom dei to. Slike tilvisingar lyt elles bli eit gjennomgåande drag som ikkje rører berre ord med jamning, og det kan derfor komma i konflikt med kravet til konsentrasjon.

Normeringa av oppslagsformer byr på ein del uløyste problem, og fleire vil sikkert melde seg etter som arbeidet går fram. Eitt som vi ser tydeleg alt no, er komposisjonsfuga i samansette ord. Her er det variasjon mellom *-a-* og *0* (*bruramarsj - brurmarsj*) eller mellom *-e-* og *0* (*enkjemann - enkmann*). Det er berre å vedgå at vi enno ikkje har funne fram til ei elegant, plassparande og samstundes lettskjøneleg løysing her.

Når det gjeld *ordninga* av oppslagsord, ser eg ikkje den systematisk-onomasiologiske som noko reelt alternativ til den alfabetisk-semasiologiske i vår samanheng. I dei fleste tilfelle vil den siste vera den mest brukarvennelege. Ein kan tenkje seg avvik frå den strengt alfabetiske ordninga stundom, ved at ein fører i lag samansetjingar med felles førstelekk (denne kan vera eit ord eller eit ord i ei viss tyding), såkalla nøsting. Like vel er det eit spørsmål om ikkje det aller mest brukarvennelege er å føre avleiingar og samansetjingar som eigne oppslag, ut frå det synet at brukaren søkjer nett det ordet som har valda samanbrot i kommunikasjonen (Cowie 1990:692).

Tilvisingar frå dialektform til oppslagsform er alt nemnt. Krysstilvisingar mellom synonym/heteronym og/eller mellom ord i same semantiske felt har vi arbeidd lite med, men eit materiale innsamla gjennom spørjelister kunne for så vidt innby til dette.

Opplysningar om uttalen vil måtte bli heller grove. Vi er redde for at for mange fonetiske spesialteikn kan verke avskrekkande på dei brukarane vi først og fremst er ute etter (jf. Hildebrandt 1986:30). Vi satsar på å greie oss så langt som råd med det vanlege alfabetet, men vi må truleg ha eigne teikn i alle fall for tjukk l og palatale konsonantar. Den ekstra vokalen, mellomvokalen mellom *æ* og *ø* som finst i mange mål i dette området, har vi førebels markert med svensk *ö*. Truleg kan ein del plass til uttaleopplysningar sparast i einskildartiklane ved å ha eit oversynskapittel som syner den viktigaste systematiske variasjonen, t.d. kvar ein har monoftongering, kvar ein har tjukk l eller r av *rd* og palatal eller dental i trykklett staving. Reglane for uttalemarkering vil sjølv sagt måtte vera dei same i hakeparentesen (dersom ein skal ha noko slikt!) som i bruksdøma.

Grammatiske opplysningar er noko som lir under innsamling etter spørjelistemetoden, og det skapar problem for ordbøker med slikt datagrunnlag. Vi har gong på gong presisert for heimelsfolka våre at dei bør opplyse om kjønn og bøyning på orda. Sjølv om somme er flinke til dette, blir det diverre jamt over nokså tynt. Medvitet om slike tilhøve er mindre enn om det reint leksikalske, og innverknad frå skriftspråk gjer også at uvissa blir stor (Witkowski 1988:61). I tilfanget vårt ser vi ofte døme på at informantane forsyner dialektord med riksspråklege endingar, som t d *-er* i fleirtal av hankjønnsord, eller at dei skriv bindevokal *-e-* i komposita som etter dialekten ikkje skal ha det. I ei folkeutgåve lyt ein vel avgrense den grammatiske informasjonen nokså mykje, og i alle høve nytte så lite som råd av grammatisk fagterminologi. Ein må i det heile vera varsam med å utvikle eit altfor avansert apparat av forkortingar, stikkord o.l. Men allmenne teikn for ordklassar, genus, verbtider osb. kjem ein ikkje utanom.

Så langt i prøveredigeringa vår har vi nytta den vanlege rekkjefølgja med formdelen før den semantiske delen. Men det spørst om vi ikkje bør tenkje på å gjera det motsett. I *Thüringerordboka* har dei gjort det, og her er også formopplysningane i mindre trykk (Spangenberg 1988:43). Også Hermann Niebaum (1986:133) går inn for ei slik rekkjefølgje.

### Utvalet av oppslagsord

Ei dialektordbok utarbeidd etter vitenskaplege prinsipp bør sjølv sagt ta med *alt* ordtilfanget som er belagt i dialekten. Det går da også fram av dei opphavlege planane for OTM at det er dette ein har tenkt seg, nemleg å gje eit oversyn over den samla ordmassen i trøndermåla. Dette kravet til å vera fullstendig lyt praktiserast innafor rimelege grenser, slik at ein t.d. kan utelata sjølv sagde og meir tilfeldige samansetjingar. Ei avgrensing kan ein også oppnå ved å stille krav om at orda skal vera "assimilerte" i dialekten. Dermed kan ein halde ute t.d. heilt nye teknisk-faglege ord.

I ei folkeutgåve må det nok veljast strengare. Her må ein leggje hovudvekta på det som er annleis, som vik av frå det riksspråklege. Ein går altså i retning av det kontrastive prinsippet, idiotikon-prinsippet om ein vil. Eit ord i vanleg riksspråkleg tyding vil hos oss i høgda bli referert stutt. Dersom det ikkje finst i andre, meir spesielle tydingar, kan det komma på tale å stengje det heilt ute. Om t.d. *hest* hadde hatt berre den vanlege dyretydinga i trøndermåla, ville vi kanskje ikkje gje det ei ordbokinnførslе i folkeutgåva i det heile (men naturlegvis i den vitenskaplege versjonen!), men sidan det også mellom anna har den meir spesielle tydinga 'stol på fele', vil ein måtte føre inn begge. Eit argument for å føre inn *hest* i vanleg tyding også om det hadde hatt berre den, er at det går inn med denne tydinga i samansetjingar som heilt klårt lyt vera med, som *hestgreie* 'hestesele', *skjuthest* 'hingst, grahest' o fl. Eit anna er at det kan gå inn med si vanlege tyding i ordtak og talemåtar som er særmerkte for dialektområdet.

For heilt allmennspråklege ord kan ein tenkje seg følgjande framgangsmåte: Den allmenne tydinga og bruken blir heilt stutt referert, medan ein framstiller særmerkt bruk meir utførleg, slik ein kan finne han i spesiell fraseologi, faste vendingar og ordtak. Ein utveg kan vera rett og slett å setje "som i riksspr" e l, og dermed føresetje skriftspråkskompetanse hos brukarane, eller vise til dei aktuelle tydingane i eit standard oppslagsverk ein kan gå ut frå at brukarane har, som t.d. *Nynorskordboka*. Det kan likevel reisast sterke innvendingar mot å føre inn tydingsopplysningar av typen "som riksspråk"; kanskje er det beint fram uforsvarleg (Lindgård Hjorth 1989:49). Etter Vidar Reinhammar (1971:41) er det utanom "sakliga

konkreter" unntaket meir enn regelen at ein har fullstendig synonymi mellom eit dialektord og eit riksspråkkord.

Handteringa av ord som viser få dialektale leksikalske særdrag, er kanskje det området der det først og fremst krevst nytenking i dialektordbøkene (Bergmann 1988:49).

Også i ei dialektordbok som legg vinn på å vera vitskapleg og fullstendig, vil det vera slik at ein legg størst vekt på det særmerkte og såleis i viss monn nærmar seg ei kontrastiv framstilling. Som Günter Bergmann (1988:49) seier, er det jo ord og tydingar som er framande for riksspråket som er dialektleksikografiens eigentlege metier.

Jamvel om ein berre skal ta med det som er særmerkt for dialekten, vil ein fort få eit omfangsproblem, i alle høve i ei folkeutgåve. Ei mogleg løysing kan vera å setje eit minimumstal på herad eit ord må vera heimla i for at det skal først opp i ordboka.

Reint *individuelle* ord er eit område der det også kan vera mogleg å avgrense. I innleiinga til *Jysk Ordbok* (Indledning XXI - XXII) er dette drøfta. Ein konkluderer her med at ein i dialektordbøker bør vera meir imøtekommande for individuelt ordforråd. Det *kan* vera restar av utdøyande ord ein har å gjera med, eller ein kan ha registrert eit ord i fødselen som kan komma til å bli meir allment akseptert etter kvart. Same kor grundig ein kjenner eit målføre, kan det stadig dukke opp ord ein ikkje var klar over fanst der. For eldre periodar registrerer ein hapax legomenon (ord som kan påvisast berre ein stad) utan å nøle.

I ei folkeutgåve kan ein nok tillata seg å vera noko meir restriktiv overfor heilt individuelle ord. Det er heilt klårt opplysningar i tilfanget vårt som verkar nokså individuelle, tilfeldige. Det kan vera ord som avvik frå det geografiske utbreiingsmønsteret ein elles kjenner, eller som er formelt umogleg i dialekten. Stundom har vi ved nærmare undersøking kunna avdekkje slike mistenkelege tilfelle som lån som heimelspersonen kan ha snappa opp av folk frå andre stader, under arbeidsopphald borte frå heimbygda osv. Innsamlings-situasjonen med spørjelister *kan* også innby til konstruerte løysingar, dersom informanten (meir el mindre umedvite) oppfattar skjemaet som ein slags eksamen der det gjeld å ha svar på så mange spørsmål som råd.

Eit problem som også gjeld utvalet av ord, er det ein kan kalle *forfattarord med tydeleg dialektpreg*. Mange forfattarar utnyttar dialektgrunnlaget sitt og er klårt merkte av det, også utanom replikkane. Typiske eksponentar for dette innafør trøndersk er Olav Duun og Inge Krokann. Ein må ha klårt for seg at slike forfattarar nyttar dialekten som eit kunstnarleg verkemiddel, og at dei ikkje har som sitt primære føremål å dokumentere dialekten. Men dette tilfanget bør så absolutt utnyttast. Dialektreplikkane kan gje gode døme på bruken av orda. Det er opplagt at ein forfattar som Duun i stor grad reproduserte talemåtar og heile setningar han hadde høyr i heimemiljøet, han hadde fullt av slike notatar. Det ville vera uheldig å avskjera seg frå slike kjelder.

Men kva gjer ein med ord som ikkje er riksspråklege, som berre finst hos forfattaren, og som ikkje er belagt i munnleg dialekttilfang? Eit konkret døme på dette er ordet *andløge* hos Inge Krokann: "Med eitt frys alt andløgje ut av honom" (*Gjenom Fonna II*: 111-112). Vi kan forsøksvis setje tydinga til 'det å vera låttsprengd, låttmild', som på vårt område er belagt særleg frå dei indre bygdene i Sør-Trøndelag (Oppdal, Rennebu, Meldal, Budal, Singsås, Haltdalen, Ålen, Røros). Det finst dessutan i Sunndalen på Nordmøre, og Ross har det frå Fosen. Også sønnafjells er det kjent fleire stader frå, både på Aust- og Vestlandet. Men substantivet *andløge* finst altså berre hos Krokann.

Av ein eller annan grunn er det ikkje teke med i *Norsk Ordbok*, enda det er der det har vorte ekserpert. Det kan neppe vera fordi det er hapax legomenon, for det manglar elles ikkje på dette i *Norsk Ordbok*. Oppdalsmålet er eit av dei bygdemåla som er aller best kartlagt

leksikografisk, først og fremst gjennom Rise 1933 og Donali 1988. Men noko *andløge*, n finst det ingen opplysningar om.

Har så Krokann laga dette ordet sjølv, eller har han høyrte det i gammalt oppdalsmål, som han kjende til botnar? Som nemnt er det farleg, sjølv med svært grundig kjennskap, å seie bestemt at eit ord ikkje finst eller ville vera mogleg å bruke i målet. For min eigen del kjem eg stadig på sporet av ord eg ikkje har visst fanst på dei stadene eg har mitt hjertespråk frå, Sunndal og Øksendal. Forfatternen kan ha hatt interesse av å blåse liv i eit ord han har likt, men som eigenleg ikkje høyrer heime i dialekten (jf. Baur 1986:81). Men det aktuelle ordet her har sterk tilknytning til det velkjende *andløg*, og eg er derfor stemt for å ta det med. I ein vitskapleg versjon bør det i alle høve vera med!

### Sider ved tydingsdelen

Også i denne delen må slankeprinsippet gjera seg sterkt gjeldande i ei folkeutgåve. Ein kan t.d. ikkje leggje opp til ei like detaljert inndeling i tydingar og ei like grundig heimling av tydingane som ein ville gjera i ei strengt vitskapleg utgåve. Ein kan heller ikkje ta sikte på ei fullstendig oppføring av kollokasjonar, syntaktiske samband osv. Også her må ein leggje størst vekt på det særmerkte, ikkje-rikskspråklege.

Redaksjonelle karakteristikkar, kjeldetilvisingar osv. lyt ein akte seg for å gje eit for sterkt kode-preg. Det spørst om dei litteraturforkortingane vi til no har nytta, faktisk er for vidløftige. Opplysningar om geografisk utbreiing trur eg derimot ein skal vera varsam med å skjera for mykje ned på eller markere for grovt. Slikt er gjerne noko av det folk spør etter.

Skiljet mellom sak og språk er eit grunnproblem i leksikologi og leksikografi, sidan den leksikalske funksjonen står svært nær røynda leksema representerer (Coseriu 1970:9). Det synest vera allmenn semje om at det bør vera plass for noko meir saksinformasjon (encyklopedisk) i dialektordbøker, mellom anna fordi dei representerer ein kultur som vil bli meir og meir fjern for ettertida. Dette problemet er nok ikkje fullt så påtrengande i Noreg som i mange andre land, men også her, som dialektane held bra på det formelle/strukturelle grunnlaget, forandrar ordtilfanget seg temmeleg fort. Gruppespråklege ordbøker vil elles i det heile vera meir saksnære enn allmennspråklege, med glidande overgangar til leksika (Reichmann 1969:95). Vi vil definitivt prøve å te oss slik at det vi gjer, kan kallast *språkbeskriving*, men gje rom for opplysningar om folketru, eldre arbeidstilhøve osv. Ei dialektordbok er ikkje berre for dei reint språkinteresserte, men også materialgrunnlag for t.d. etnologar, folkeminnegranskarar og historikarar.

Ved utforminga av definisjonane vil vi prøve å halde oss innfor eit vanleg kvardagsspråk så langt råd. Vi vil heller ikkje gå av vegen for innslag av ein noko meir pratande, pludrande stil enn den tradisjonelt strenge ordboksprosaen, så lenge det ikkje tek overdriven plass. Ei anna side ved sjølvdefineringa i dialektordbøker er at ein her, i alle høve dersom ein følgjer eit kontrastivt utvalsprinsipp, i stor mon kan bruke synonymdefinisjonar og dermed spare plass.

Tilhøvet mellom tydingar og bruksdøme fortener nærare vurdering. Bergmann (1988:53) går på eit vitskapleg grunnlag inn for å leggje større vekt på bruk enn tyding i dialektordbøker, og ta med rikeleg med gode dialektdøme også på kjente tydingar. Ein tek da eit steg i pragmatisk/diskursiv lei, kjem nærare brukssituasjonen, og møter dermed den innvinginga som ofte har vore retta mot ordforskinga, at ho ikkje tek nok omsyn til moderne kommunikative/pragmatiske tilnæringsmåtar. Også vi er opne for ei slik omvurdering i vår

folkeutgåve, men da ut frå at gode bruksdøme vil vera artig og interessant lesnad for den brukargruppa vi er ute etter.

Utan at vi har fått arbeidd særleg mykje med det til no, er det klårt at gode illustrasjonar vil friske opp i ei folkeleg dialektordbok. Ein kan tenkje seg både utbreiingskart og teikningar av sentrale reiskapar som t d *sleden* og *hesteselen*. Desse tek rett nok i seg sjølv stort rom, men kan på den andre sida gje høve til å spare noko på ordbruken i definisjonane.

### Den historiske dimensjonen

Kjeldetilfanget blir tunnare di lenger attover ein kjem i tid, men noko finst. Alt hos Aasen og Ross er opplysningane no hundre år gamle og meir, og det kan vera interessant å sjå kva som er gått ut og kva som er ført vidare til våre dagar av ordtilfanget dei registrerte. For trøndermåla sin del finst det også ein god del eldre tekster, for perioden 1706-1856 samla i Dalen/Hagland 1985. Det er teke vare på nokre få ordsamlingar frå 1700-talet. I avisa *Næmingen* finst barnetekster på dialekt frå rundt siste hundreårsskifte.

Andre moglege kjelder til ordtilfanget i eldre tider er gamle skifte, inventarielister m m. Utnyttinga av dette blir likevel meir tilfeldig, alt etter hva som kan vera prenta i lokalhistorisk litteratur, kva vi har fått av lokalhistorikarar vi har vore i kontakt med osb. Tid og ressursar tillet inga fullstendig ekserpering her, og det same gjeld aviser.

Vi vil likevel utnytte det historiske materialet så godt vi kan, og ser det som viktig, i alle høve i den vitskaplege delen av prosjektet.

Også i folkeutgåva meiner eg det bør gjevast rom for etymologiske opplysningar. Folk spør ofte etter kva ord kjem av, og dei bør finne svar på det, også på ord der ein kan finne etymologien i etymologiske ordbøker eller andre oppslagsverk. Ein kan ikkje utan vidare rekne med at brukarane har tilgang til slikt. Sjølve lemmatiseringa kan vera resultat av etymologisk granskning, og oppslagsforma kan seie ein god del om opphavet til ordet.

### Publisering av delar av tilfanget

Når innsamlinga er gjort ved spørjelister om avgrensa emne, kan det liggje an til å publisere ordtilfang frå bestemte saksfelt. Dette har til no gjeve seg utslag i ei trøndersk skjellsordbok (Jenstad 1991), basert først og fremst på ei spørjeliste om karakteristikk av personar. Vi ser ikkje bort frå at denne boka vil fenge og bli ein appetittvekkjar, ein forsmak på kva ein kan vente seg når heile tilfanget ein gong kjem ut.

### Folkeutgåve - eller ikkje?

Som det har gått fram, er vår folkeutgåve enno mykjegodt på tenkje- og prøvestadiet. Men om vi får bestemt ein høveleg framgangsmåte, ein framstillingsmal, skulle arbeidet kunne gjerast på rimeleg tid. Det er like vel gode argument *mot* ei slik oppdeling: Bør ein ikkje rekne med at ei trøndersk dialektordbok er noko som i alle høve kjem til å bli laga berre ein gong? Bør ein ikkje greie å kombinere det vitskapleg forsvarlege med det allment forståelege? Vil ikkje ordboka i alle høve først og fremst bli eitt reiskap for forskarar (jf Grosse 1988:28)? Vi er også kjende med at ein i *Jysk Ordbok* har gjeve opp tanken om ei folkeutgåve. Hovudgrunnen til at vi tenkjer oss ei deling, er *tidsaspekter*: Ei fullstendig vitskapleg

utarbeiding vil trekkje ut i tid, og vi følger forplikting og forventningspress om å få gjeve ut eit resultat relativt raskt.

### Litteratur

- Baur, Gerhard W. 1986. Quellen und Corpora. Zur Materialbasis deutschsprachlicher Dialektwörterbücher. I: Hans Friebertshäuser, utg. *Lexikographie der Dialekte. Beiträge zu Geschichte, Theorie und Praxis*: 75-91. Tübingen.
- Bergmann, Günter. 1988. Zur Darstellung der Semantik im OSWb. I: *Dialekt-Lexikographie. Berichte und Analysen zur Arbeit an Dialektwörterbüchern*: 47-57. Jena.
- Coseriu, Eugenio. 1970. *Einführung in die strukturelle Betrachtung des Wortschatzes*. Tübingen.
- Cowie, A.P. 1990. Language as Words. Lexicography. I: N.E. Collinge, red. *An encyclopedia of Language*: 671-700. London/New York.
- Dalen, Arnold og Hagland, Jan Ragnar. 1985. "I det meest upolerede Bondesprog". *Tekster på trøndermål 1706-1856*. Oslo - Bergen - Stavanger - Tromsø.
- Donali, Ingeborg. 1988. *Oppdaling. Ord og uttrykk (Oppdalsboka III)*. Trondheim.
- Grosse, Rudolf. 1988. Einige Überlegungen zu den Grundlagen der Dialektlexikographie. I: *Dialekt-Lexikographie. Berichte und Analysen zur Arbeit an Dialektwörterbüchern*: 27-34. Jena.
- Hildebrandt, Reiner. 1986. Dialektologie und Dialektlexikographie. Zum Stellenwert einer Subdisziplin. I: Hans Friebertshäuser, utg. *Lexikographie der Dialekte. Beiträge zu Geschichte, Theorie und Praxis*: 21-33. Tübingen.
- Jenstad, Tor Erik. 1984. Arbeidet med trønderordboka. I: *Motskrift* 2: 25-43. Trondheim.
- Jenstad, Tor Erik. 1989. Noko om arbeidet med Trønderordboka. I: Olaf Almenningen & Oddrun Grønvik, red. *Ord og mål. Festskrift til Magne Rommetveit 4. oktober 1988*: 120-127. Oslo.
- Jenstad, Tor Erik. 1991. *Skjellsordboka*. Trondheim.
- Jysk Ordbog*. Bind I. Hæfte I. Aarhus 1970.
- Killingbergtrø, Laurits. 1983. Ord og uttrykk frå dialekten. I: Asbjørn Steen, red. *Leksvik Bygdabok. Kulturbind II*: 67-143. Leksvik.
- Lindegård Hjorth, Poul. 1989. Leksikografi. I: Bente Holmberg, Britta Olrik Fredriksen & Hanne Ruus, red. *Forskningsprofiler*: 27-63. København.
- Niebaum, Hermann. 1986. Lemma und Interpretament. Zur Problematik der Artikelgestaltung in Dialektwörterbüchern. I: Hans Friebertshäuser, utg. *Lexikographie der Dialekte. Beiträge zu Geschichte, Theorie und Praxis*: 125-143. Tübingen.
- Reichmann, Oskar. 1969. *Deutsche Wortforschung*. Stuttgart.
- Reinhammar, Vidar. 1971. Gustav Adolfs Akademiens dialektordbok. I: *Saga och sed*: 29-53. Uppsala.
- Rise, Ola J. 1933. *Opdalsmål. Serord og talevendingar*. Svorkmo.
- Spangenberg, Karl: Stichwortansatz und Artikelaufbau im Dialektwörterbuch. I: *Dialekt-Lexikographie. Berichte und Analysen zur Arbeit an Dialektwörterbüchern*: 35-46. Jena.
- Witkowski, Teodolius. 1988. Grammatische Angaben im Dialektwörterbuch. I: *Dialekt-Lexikographie. Berichte und Analysen zur Arbeit an Dialektwörterbüchern*: 58-65. Jena.

Martin Gellerstam, Sven-Göran Malmgren og Maria Toporowska-Gronostaj

## Inlärningslexikon

Uppsatsen innehåller dels en allmän diskussion om inlärningsordböcker (=learner's dictionaries), dels några förslag angående utformningen av tre typer av information i sådana ordböcker: definitionerna, kontextuell information (valensuppgifter) och information om kollokationer. Bl. a. diskuteras a) COBUILD-inspirerade definitioner, där den syntaktiska utbytbarheten uppges b) valensangivelser av såväl implicit som explicit slag c) anordning av kollokationer på grundval av den semantiska relationen mellan de båda komponenterna i kollokationen.

### 1. Vad är ett inlärningslexikon?

I svenskan saknas den typ av större specialinriktade lexikon som i engelskspråkiga länder brukar gå under namnet "learner's dictionaries". Det finns visserligen ansatser till sådana lexikon också i svenskan, t.ex. underlaget till de tvåspråkiga ordböckerna inom det s.k. LEXIN-projektet, publicerat separat under titeln *Svenska ord med uttal och förklaringar* (1984). Det finns också åtskilligt att hämta för den som kan lite mer svenska i *Svensk ordbok* (1986), även om den ordboken i första hand är inriktad på svenskers behov.

Hur ska då ett "learner's dictionary" (här i fortsättningen kallat *inlärningslexikon*) se ut? Går vi till England - där termen myntats - så ser vi att man normalt menar en enspråkig ordbok, för såväl analys (passiv användning, "decoding") som produktion (aktiv användning, "encoding") och med bl.a. explicita grammatiska beskrivningar av ords kontextuella uppträdande. Dessa enspråkiga inlärningslexikon (COBUILD, LDOCE m.fl.) anser sig klara av konststycket att på en gång vända sig till den som lär sig modersmålet och den som lär sig ett andraspråk, den som vill slå upp ett ord i en engelsk text och den som vill formulera sig på engelska.

Samtidigt pågår sedan länge en diskussion bland lexikografer om hur ett hjälpmedel för andraspråksinläring optimalt bör se ut: ska det vara enspråkigt eller tvåspråkigt, ska man skilja på aktivt och passivt osv. Många efterlyser också en mer sammanhängande syn på lexikalisk inläring där lexikonet placeras in som en del i hela den process som inläringen av ett andraspråk utgör. Ett inlärningslexikon kan inte komma in överhuvudtaget förrän man har lärt sig vissa grunder i ett andraspråk och förmår jämföra de två språkens strukturer. Det handlar också om graden av specialisering av lexikon: mot de lexikografer som så att säga bakar in alla användningar i ett enda lexikon (typ COBUILD) står de som vill specialisera lexikon för olika användare. Som ett exempel på ett typiskt aktivt speciallexikon kan nämnas Benson & Benson & Ilsons *The BBI Combinatory Dictionary of English* (hädanefter *BBI*). Till de rent språkliga argumenten i frågan om inlärningslexikon kommer naturligtvis också förlagsmässiga argument: det är mer tilltalande att göra ett enda lexikon för en stor grupp användare än att göra flera mer specialiserade för olika typer av användare.

## 2. Skrivet eller datalagrat lexikon?

Det kan vara nyttigt att vidga perspektivet ytterligare och diskutera det skrivna lexikonets roll överhuvudtaget. I en uppsats från 1984 med titeln "The Dictionary of the Future and the Future of the Dictionary" argumenterar Martin Kay för en mer flexibel syn på lexikonet, nämligen det datalagrade lexikonet. Han vill se det som en källa till kunskap på olika nivåer, han suddar ut gränsen mellan lexikon och text (hur mycken tid har inte lexikonredaktörer ägnat åt att våga exemplen på guldvåg!), han visar hur ingången till lexikonet kan bli lättare (man behöver inte ens kunna stava rätt för att komma rätt) osv.

Tiden har hunnit ikapp många av Martin Kays konkreta exempel. Idag är en hel del lexikoninformation, särskilt kanske den som krävs för produktion av text, en integrerad del av ordbehandlingsprogram: vi kontrollerar stavning, letar efter synonymer och vandrar ut i betydelsefält. Ja, vi har t.o.m. vad företagen kallar "style checkers" som stöd i vår skrivarmöda.

Utveckling av vad man kan kalla skrivhjälpmedel finns också att hämta inom den gren av översättning som kallas CAT (=computer-aided translation). Här rör det sig om hjälp med lexikonslagning, kontroll av konstruktioner och t.o.m. automatisk översättning av vissa ord. Hjälpmöden är tillsvidare mest inriktade på hjälp till professionella översättare - men principiellt är hjälpmedlet en implementering av kunskapen i ett aktivt lexikon för andraspråksinlärare.

En hel del livligt debatterade frågor när det gäller utformningen av inlärningslexikon skulle med fördel kunna ses i ljuset av dessa nya hjälpmedel. Ett exempel: en viktig ingrediens i aktiva lexikon för andraspråksinläring är uppgifter om ords grammatiska kontext (se t.ex. LDOCE och COBUILD). Hur dessa uppgifter ska anges är en populär stridsfråga: ska man ha en explicit men svårbegriplig notation eller en implicit och lättbegriplig? Eller ska man undvika sådan information överhuvudtaget med hänvisning till att ingen ändå använder den? I ett datalagrat lexikon blir denna och andra frågor inte så viktiga: här som i andra sammanhang kan man tänka sig olika grader av exakthet och utförlighet. Kanske bör man dessutom koppla den grammatiska beskrivningen till fler exempel, man vandrar helt enkelt vid behov ut i texter eller till systematiskt sammanhållna klasser av grammatiska företeelser, till en grammatik baserad på lexikonet om man så vill. Samma sak gäller exemplifieringen av ett lexem (ett numererat betydelsemoment av ett flertydigt ord) med en viss definition: man går ut i texter och hämtar så många exempel man anser sig behöva.

Men är inte hela det här resonemanget tämligen orealistiskt? Vem vill använda ett datalagrat lexikon för att lära sig ett språk och vem ska förresten bygga upp ett sådant här lexikon, eller datalagrat språkinlärningsinstrument?

Datalagrade lexikon är redan ibland oss, med eller vanligen utan lexikografers medverkan. Vi har redan idag små relativt billiga "språkdosor" i omlopp som innehåller hundratals tusen ord i snabb access. Vad vi väntar på är större teckenruta (display), mer varierad information och lägre pris. Dessutom sparar vi ett antal hektar skog, ett argument som små småningom kommer att bli allt viktigare.

När det gäller den andra frågan tror vi att det kan vara nyttigt att påminna sig vilka olika specialkunskaper (inte sällan förkroppsligade i en person) som är inblandade i ordboksarbete: lingvisten/särspråksexperten som beskriver språkets struktur, lexikografen som samlar ihop kunskapen från olika håll, definierar orden och beskriver dem i en konsekvent och sammanhängande form, språkpedagogen som anpassar ordbokstexten till en bestämd användargrupp och systemutvecklaren som strukturerar den information som når läsaren. I utvecklingen av skrivna lexikon finns en tendens att lexikografen onödigt mycket ikläder sig



rollen av pedagog, förenklar, skär ned exempel och anpassar lexikokontexten till den presumtive användaren. I ett datalagrat lexikon kan lexikografen mer inrikta sig på sin roll att ge en varierad bild av den språkliga verkligheten utan ängsliga sidoblickar på vad formatet tål. Samtidigt måste han dock veta vad *läsaren* tål: datatekniken avlastar inte lexikografen mödan att beskriva och definiera betydelser och att illustrera analysen med goda exempel. Sedan kommer det an på lexikografen - i samarbete med skickliga konstruktörer av gränssnitt - att leda läsaren vid handen in i den lexikaliska kunskapen.

Det bör vara en utmaning inte minst för oss lexikografer att knyta ny och annorlunda strukturerad information till det som vi traditionellt kallar lexikon och som i detta perspektivet snarast är ett lexikaliskt språkinlärningsinstrument. Åtminstone bör inte ordbehandlings-tekniker ensam få definiera vad som är nyttig lexikalisk kunskap.

### 3. Hur ska ett inlärningslexikon se ut?

I det följande ska vi ta upp några typer av lexikonuppgifter och deras utformning i ett inlärningslexikon: det gäller formen för definitioner, återgivningen av kontextuella fakta av valenstyp och, slutligen, uppgifter om kollokationer. Huruvida alla dessa uppgifter kan fogas in i en och samma (tryckta) inlärningsordbok är dock en fråga som vi måste lämna öppen. Vi börjar med definitionerna.

#### 3.1 Definitioner i COBUILD-format

Lexicography is an unnatural occupation. It consists in tearing words from their mother context and setting them in rows ... Half of the the lexicographer's labor is spent repairing this damage to an infinitude of natural connections that every word in any language contracts with every other word. (Bolinger 1985).

Definitioner i enspråkiga inlärningslexikon - antingen de inriktas på produktion eller avkodning - kräver en grundläggande kunskap i målspråket och en viss insikt i hur definitioner i allmänhet är utformade. Det finns här en tradition inom engelsk lexikografi att göra definitioner med hjälp av ett någorlunda enkelt, grundläggande ordförråd (se t.ex. LDOCE). Men det är först Collins COBUILD som har tagit steget över till en typ av definition som inte bara använder enkla ord utan vars struktur är helt annorlunda.

Definitionstypen kan illustreras genom exemplet *öppna* i Figur 1 (se s. 350). Vid första anblicken ser det ut som den typen av oreflekterade definitioner som barn använder (t.ex. "Vad är kärlek" - "Det är när man tycker om någon"). Och barns definitioner - liksom COBUILD:s - har den egenskapen att de tar med aktörerna runt det ord som skall förklaras. Tar man med dessa aktörer eller medspelare så slipper man att rycka orden från deras "mother context" i lika stor utsträckning som vad fallet är i traditionella definitioner. Därigenom vinner man två saker: större enkelhet och mindre behov av metaspråklig information. Den större enkelheten har att göra med att undertryckande av information (t.ex. undertryckande av subjekt och objekt) leder till abstraktion. Till undertryckandet av information kan man också föra det välkända lexikografiska kravet på "syntaktisk utbytbart". Det minskade behovet av metaspråklig information handlar om att man slipper många preciseringar som "med avseende på", "om" osv.

Figur 1. Utkast till ordboksartikeln *öppna*

**öppna** /2öp:na/ *öppnade öppnat* verb

**1 OPEN** Om man öppnar något, t.ex. en dörr, så ändrar man dess position eller låser upp den så att luft, ljus, saker eller människor kan passera genom den.

<A öppnar x>

1.1 konkret en dörr, ett fönster, ett lås, en burk, en flaska, en bok, ett brev, en tidning, en hand, ett öga, en famn

*Kalle öppnade genast dörren när det ringde* N V N

*Mormor öppnade (grinden) och släppte in oss* N V (N)

*Tjuven öppnade låset med en skruvmejsel* N V N med N

1.2 abstrakt en dörr, ett hem, en väg

*Präster öppnade kyrkans dörrar för flyktingar* N V N för N

*Forskningen öppnar vägen till en ny framtid* N V N

1.3 idiom *öppna munnen* ("börja tala"); *öppna sitt hjärta (för någon)* ("berätta sina innersta hemligheter")

**2 OPEN** Om någon eller något öppnar sig så inträder ett nytt läge eller tillstånd.

<A/x öppnar sig>

2.1 konkret en blomma, en dal

*Blommorna öppnade sig mot solen* N Va Nr mot N

*Dalen öppnar sig mellan de branta bergväggarna* N Va Nr

2.2 abstrakt en möjlighet

*Nya möjligheter har öppnat sig för henne* N Va Nr för N

**3 OPEN UP** Om en affär eller annan offentlig lokal

öppnar vid en viss tid så låser man upp dörrarna

så att folk kan börja arbeta och besökare kan

utnyttja lokalen.

<x öppnar + TID>

en affär, ett kontor, en kyrka, ett museum

*Affären öppnar kl. 9* N Va A

**4 INAUGURATE, START** Om man öppnar en affär eller ett kontor så påbörjar man en ny (affärs)verksamhet

<A öppnar x>

*Tandläkaren öppnade praktik i Kungsbacka* N V N

idiom *öppna eget* ("öppna egen affärsverksamhet")

**5 BEGIN** Om man öppnar något så börjar man utföra något

*Löparen öppnade löst* ("började i måttligt tempo")

*men ökade efter den första milen* N V A

*Min medspelare öppnade försiktigt med 1 spader* N V A

idiom *öppna eld* ("börja skjuta"); *öppna korrespondens* ("börja brevväxla"); *öppna konto* ("börja spara pengar i en bank")

Exemplets definition av *öppna* kan jämföras med *Svensk ordboks* genomtänkta men mer abstrakta definition (här enbart det första lexemet och med endast ett exempel på varje underbetydelse; det som står inom tecknet / är definitionskommentarer).

- bringa till öppet (eller öppnare) läge  
/med avseende på spärrande anordning/ *öppna ett fönster*
- /även med avseende på utrymme etc./ bringa till ett öppet eller öppnare tillstånd *öppna en konservburk*
- /även med avseende på något som har tillsluten form; ibland ifråga om skenbar förändring i samband med någons förflyttning/ *öppna sin famn för någon*
- /även i liknelser eller överfört, särskilt med tonvikt på avlägsnande av hinder för utveckling eller uppmärksamhet/ *öppna portarna för en ny generation konstnärer*

Styrkan med definitioner av *Svensk ordboks* typ är att man utgår från en central betydelse och sedan ger andra betydelser som utvidgningar, inskränkningar, metaforer etc. utifrån denna grundbetydelse. Detta ger en generalitet i beskrivningen som är sådan att eventuellt nya betydelser normalt kan infogas under respektive lexems huvudbetydelse. Det pris man måste betala är enkelhet i definitionerna. COBUILD-definitionerna har inte på samma sätt ett tydligt centrum från vilket andra betydelser utgår - exempelvis använder man inte den metaspråkliga beteckningen "överförd betydelse" utan ger den istället i klartext utan explicit referens till grundbetydelsen. Definitionerna är enkla men ändå precisa och kompletteras dessutom av explicit grammatisk information. Men den precision och enkelhet man åstadkommer vid beskrivningen av enskilda betydelser har sitt pris i form av svåröverskådlighet (som inte blir mindre genom att uppslagsordet utgörs av ordets grafiska form och olika ordklasser på detta sätt samsas under ett huvudord).

Sammantaget har ändå COBUILD-definitionerna sådana pedagogiska fördelar att man allvarligt måste fundera på om de inte skall finnas med i en inlärningsordbok.

### 3.2 Explicit grammatisk beskrivning

Ett av syftena med ett inlärningslexikon är att förankra definitionens semantiska innehåll i språkets grammatiska system. Någon form av kontextuella uppgifter, valensuppgifter, är därför nödvändiga i varje inlärningsordbok. Av existerande svenska ordböcker är det främst *Svenska ord* som ger explicita valensuppgifter, med hjälp av ett rätt avancerat kodsysteem. Det system som vi här vill föreslå (se åter artikeln *öppna*) bygger på systemet i *Svenska ord* men har även tagit upp idéer från nyare valensteoretisk litteratur (Allerton 1986, Nikula 1986, Somers 1987, Gellerstam 1988, Toporowska-Gronostaj 1991).

En bärande tanke i vårt förslag är att ordboken bör ta hänsyn till olika användares behov. I valet mellan att ge explicit men svårbegriplig och lättbegriplig men implicit information (jfr ovan) väljer vi att ge båda. Vårt förslag innebär en uppläggning i två steg. Det första steget innebär att man tillhandahåller ett urval exempel, där de mest typiska

konstruktionerna ges. I det andra steget ges koder som sammanfattar de aktuella konstruktionerna på ett mer formaliserat sätt.

Följande grammatiska kategorier förefaller vara nödvändiga vid beskrivningen av verbvalensramar:

N	nominalfras
V	verbfras
A	adverbfras
Adj	adjektiv
I	infinitiv
S	sats

De ytterligare koder som använts i artikelns syntaktiska verbvalensramar syftar till att fånga semantisk-syntaktiska restriktioner på konstruktionens produktivitet. Som ett exempel på syntaktisk produktivitet kan nämnas det transitiva verbets benägenhet att förekomma i passiva konstruktioner. När det gäller passiveringen av verbet *öppna* så framgår det av valenskoderna att lexem 1 och 4 tillåter passivering men inte lexem 2 och 3. Koden "a" hänför sig till aktivum.

Som ett ytterligare exempel på en restriktion av semantisk-syntaktisk natur kan nämnas "Nr" där "r" begränsar nomenklassen till reflexiva pronomina.

Av intresse för valensen är att notera de fall där det råder diskrepans mellan semantisk valens, dvs. semantiska roller (ungefär = kasusroller av Fillmore-typ), och syntaktisk valens, vilket illustreras nedan:

- |      |                       |
|------|-----------------------|
| (1)  | Jan bor i Göteborg    |
| (1') | Jan bor lyxigt        |
| (2)  | Jan vistas i Göteborg |
| (2') | *Jan vistas lyxigt    |

Av dessa exempel framgår att för verbet *bo* bör ges två valens-koder, men för verbet *vistas* bara en.

Explicit anges i valensramarna prepositioner som förekommer i de syntaktiskt och/eller semantiskt relevanta argumenten, vilket inte bara styr det rätta lexikaliska valet utan även antyder de semantiska roller som associeras med den aktuella betydelsen.

Utöver de koder för de enskilda exemplen som står till höger om satserna ges en generell valensstruktur omedelbart efter definitionen. Den koden sammanfattar grundläggande semantiska och syntaktiska fakta för korrekt användning av ordet. A, B, C står för personer, x, y och z för saker eller abstrakta företeelser.

Sammanfattningsvis kan sägas att den föreslagna ordboksartikelns utformning med tydligt markerade sektioner för de formella och innehållsliga delarna bör underlätta för användaren att snabbt orientera sig i artikeln. Han förses med uppgifter om uppslagsordets uttal, ordklass och böjningsformer. För den innehållsliga delen föreslås dessutom information om det semantiska fält som aktualiseras i definitionen. Denna information ges på engelska för att vägleda användaren på ett språk som han möjligtvis behärskar bättre än svenska. Explicit hänvisning till manifestationer av de semantiska rollerna utgör ett ytterligare stöd för användaren, som kan sakna eller vara osäker på selektionsrestriktionerna för uppslagsordet ifråga.

### 3.3 Kollokationer; kombinationsordböcker

På senare år har termen *kollokation* kommit alltmer i förgrunden, inte minst i debatten om inlärningslexikon. Med denna term avses här, i anslutning till Hausmann (1985), en pregnant tvåordskombination (obestämda artiklar o.d. oräknade) av typen *fatta beslut, spela in en film, en inbiten unghäls, en kronisk alkoholist*. Sådana tvåordskombinationer måste anses ingå i språkbrukarnas abstrakta lexikon i den meningen, att dessa för att kunna anses ha tillräckliga kunskaper om exempelvis ordet *beslut* nödvändigtvis måste behärska frasen *fatta beslut*. Eftersom verbet *fatta* i denna användning är helt oförutsägbart, måste frasen *fatta beslut* anses lexikaliserad, dvs. lagrad i språkbrukarnas inre lexikon. I saussureska termer kan man med Hausmann uttrycka det så, att pregnanta fraser som *fatta beslut* tillhör *langue* i motsats till banala fraser som *ett oklokt beslut*, som enbart tillhör *parole*.

En kollokation består sålunda av två komponenter, som kan kallas *bas* och *kollokator*. (Hausmanns terminologi, som vi här lånar, är mer utarbetad än Svenséns (1987, s. 93) och därför att föredra.) I exempelvis frasen *fatta beslut* är *beslut* *bas* och *fatta* *kollokator*, i frasen *en inbiten unghäls* är *unghäls* *bas* och *inbiten* *kollokator*. Det implicita påstående om en hierarkisk struktur som ligger i termen *bas* kan enklast motiveras på följande sätt. Vi tänker oss en textproducent med annat modersmål än svenska, som vill uttrycka idéinnehållet 'fatta ett beslut'. Det är fullt tänkbart att denne textproducent känner till ordet *beslut* i isolering, men att han är osäker om det kollokerande verbet *fatta*. Det motsatta fallet, att han skulle känna till *fatta* i just denna användning men inte *beslut*, är nästintill otänkbart. Från textproducentens synpunkt är ordet *beslut* därför primärt. (Resonemanget är naturligtvis också tillämpligt om vår textproducent har svenska som modersmål.)

Men nu kommer det från både teoretisk och praktisk synpunkt intressanta problemen: hur ska man utforma en huvudsakligen *aktiv* enspråkig kollokationsordbok av inläringstyp? Hur ska vår icke-svenske textproducent hitta fram till den riktiga kollokationen *fatta beslut* med hjälp av existerande enspråkiga svenska ordböcker? Han får knappast tillräcklig hjälp ens av den ordbok som mest liknar en kollokationsordbok, *Svensk handordbok*; även om frasen *fatta beslut* naturligtvis tas upp i denna ordbok kan han inte utan vidare veta att just den frasen betyder det han vill uttrycka. På detta problem erbjuder den typ av ordböcker som ibland kallas "explanatory combinatorial dictionaries" (låt oss för enkelhets skull kalla dem *kombinationsordböcker*) något av en lösning - åtminstone i många fall. Man kan säga att dessa ordböcker bygger på iakttagelsen att det finns en viss uppsättning semantiska relationer mellan *bas* och *kollokator* som återkommer i ett stort antal kollokationer. Exempelvis förekommer relationen 'ge upphov till, komma att existera' såväl i kollokationen *fatta beslut* som i kollokationen *spela in en film*. Annorlunda uttryckt är *fatta* i *fatta beslut* och *spela in* i *spela in en film* just i dessa användningar mer eller mindre synonyma. I en kombinationsordbok disponerar man nu artiklarna på grundval av de semantiska relationer som etableras i ordbokens inledning; användaren vet sålunda exakt var i artikeln han finner exempelvis den kollokation som svarar mot relationen 'ge upphov till'. Den här aspekten på kollokationer har främst studerats av sovjetiska lingvister inom ramen för den s.k. Meaning-Text-theory (MTT; se t.ex. Melčuk & Zholkovsky 1984). De semantiska relationer (t.ex. 'ge upphov till') som just diskuterats kallas inom denna teori *lexical functions (LF)*, en term som lätt låter sig försvenskas till *lexikala funktioner*.

Med all respekt för det ytterst viktiga bidrag MTT givit till kollokationsforskningen måste det ändå konstateras att det ännu återstår att skriva en ens tillnärmelsevis fullständig kombinationsordbok över något språk, även ryska. Men MTT i förenklad och populariserad

form ligger uppenbarligen till grund för den tidigare nämnda BBI-ordboken. I BBI etableras sju lexikala funktioner, bl.a. den här diskuterade 'ge upphov till'. Motsvarande kollokationer kallas *lexikala kollokationer* (i denna term ligger att både basen och kollokatorn tillhör de öppna ordklasserna). Även ett stort antal *grammatiska kollokationer* (kollokationer där ett av de ingående orden tillhör någon av de slutna ordklasserna) urskiljs i BBI, men i detta avseende är ordboken knappast originell - det rör sig här helt enkelt om vad som i föregående avsnitt kallats valensinformation. Det - i anglosaxisk och kanske över huvudtaget i västerländsk lexikografi - unika med BBI är dess behandling av de lexikala kollokationerna. Låt oss helt kort se på de sju typerna (vi tillåter oss att något förkorta, i något fall "parafrasera", definitionerna av L1-L7):

*Exempel*

L1 Creation and/or activation	compose music; apply a principle
L2 Eradication and/or nullification	break a code
L3 (Typical) adjective + noun	a chronic alcoholic; a rough estimate
L4 Noun + (typical) verb	adjectives modify; bombs explode
L5 The <i>unit</i> associated with the noun	
a. the larger unit	a colony of bees
b. the specific unit of something larger	an act of violence
L6 (Typical) adverb + adjective	deeply absorbed; keenly aware
L7 Verb + (typical) adverb	affect deeply; argue heatedly

Allmänt kan det sägas att BBI-ordboken har en rätt generös tolkning av de lexikala funktionernas innehåll. Strängt taget innebär också endast L1, L2, L4 och L5 någorlunda konstanta semantiska relationer mellan bas och kollokator; de övriga är betydligt mer lösliga. L3, L6 och L7 är uppenbart inspirerade av den kanske mest berömda lexikala funktionen i MT-teorin, den som där kallas MAGN och som har den ungefärliga innebörden 'stort omfång, stor utsträckning' - men L3, L6 och L7 är betydligt mindre väldefinierade än så. Inget ont i det, men kanske hade det varit bättre att skilja tydligare mellan "hårda" (väldefinierade) och "mjuka" lexikala funktioner. En sak som kan förvåna är att man i själva ordboken delvis frångår den numrering, bl.a. L1-L7, som upprättats i den teoretiska inledningen, något som måste minska ordbokens användbarhet för icke-engelsktalande. Om exempelvis L1 och L3 men inte L2 är tillämpliga på ett visst ord tas L3-kollokationen sålunda upp som moment 2. Men hur ska en naiv användare kunna veta att det här är fråga om L3 och inte L2 eller L4? Med andra ord: en av de absoluta huvudpoängerna med kombinationsordböcker går faktiskt delvis förlorad.

Vi avslutar med en BBI-inspirerad lista över möjliga lexikala funktioner/kollokationer med substantiv (N) som bas i svenskan. Kollokatorerna kan vara substantiv, verb (V) eller adjektiv (A). Fastän det egentligen spränger kollokationsbegreppet såsom det definierats här accepteras också morfem i sammansättningar där N ingår (och i exemplet *bok* nedan t.o.m. sammansättningar utan morfematisk släktskap med basen *bok*):

	<i>Gram. struktur</i>	<i>Semantisk relation</i>	<i>Exempel</i>
L1	V + N	'ge upphov till, komma att existera'	spela in en film, fatta beslut
L2	V + N	'aktivera' resp. motsatsen	starta/stänga av motorn; tända/släcka lampan
L3	V + N	'det man typiskt gör med N'	köra/åka bil; dricka te
L4	N + N el. sms.	'samling, grupp av N'	en flock vargar; en bisvärm
L5	N + N el. sms.	'viktig, framträdande del av N'	(stad) stadsdel; (sjukhus) avdelning
L6	N + V	'typisk aktivitet e.d. av N'	hundar skäller; vägen går/leder till...
L7	A + N	'typisk egenskap hos N'	starkt/svagt kaffe; en blodig biff
L8	V + N	blandat	
L9	mest sms.	blandat	

Parallellerna med BBI:s system är uppenbara. Dock har en viss ytterligare indelning genomförts; exempelvis är BBI:s L1 uppdelad i två grupper. En tanke är också att de mest väldefinierade kollokationstyperna bör samlas i början och att typer med mer blandat innehåll bör tas in i grupper av karaktären "övrigt" mot slutet. Med tanke på icke-svenska användare kan det semantiska notationssystemet behöva förfinas på många sätt. I grupp 2 kan man t. ex. markera 'aktivering' (*starta, tända*) med + e. d. och motsatsen (*stänga av, släcka*) med -. Man kan vidare tänka sig undergrupper (i princip av typ konkret/abstrakt) till L1. Detta framgår också av följande utkast till en artikel *bok* i en tänkt svensk kombinationsordbok:

*bok* ('trycksak')

- L1 /abstrakt/ skriva, författa en bok; /konkret/ ge ut, publicera, trycka en bok
- L2 + slå upp, öppna en bok; - stänga, slå igen en bok, /expressivt/ smälla igen en bok
- L3 läsa (igenom, ut) en bok; bläddra i en bok; slå i en bok
- L4 en samling böcker; boksamling; bokhandel; bibliotek; antikvariat
- L5 /konkret/ bokens (10) kapitel; bokens (200) sidor; bokens omslag; /abstrakt/ bokens förord, inledning; bokens titel
- L6 boken kommer ut; boken handlar om (behandlar, avhandlar) ngt
- L7 /konkret/ en häftad bok, en inbunden bok; en tjock bok, en tunn bok; /abstrakt/ en intressant bok, en spännande bok
- L8 översätta, recensera, korrekturläsa en bok; tillägna, dedicera en bok till ngn; det står i boken att...
- L9 barnbok, pojkbok, flickbok, vuxenbok; äventyrsbok, sagobok, bilderbok; psalmbok, andaktsbok; dagbok, anteckningsbok; lärobok, läsebok, räkebok, övningsbok; handbok, kokbok; ordbok

Utöver de lexikala kollokationerna bör i detta fall även åtminstone två grammatiska kollokationer tas upp: *en bok om ngt*, *en bok av ngn*.

En sista fråga (jfr ovan) är om kollokationerna ska ges "rena" eller förses med extra kontext. BBI-ordboken är ganska generös med kontext; här har det givits mer sparsamt, egentligen endast under L5. Naturligtvis ökas (den icke-svenske) användarens chanser att förstå kollokationens innehåll av lämplig kontext, men samtidigt kan kontexten försvåra identifikationen av själva kollokationen. En möjlighet är att laborera med någon form av parentessystem som här under L5.

### Litteratur

- Allerton, D.J. 1982. *Valency and the English verb*. London.
- Benson, Morton, Benson, Evelyn & Ilson, Robert. 1985. *The BBI Combinatory Dictionary of English*. New York.
- Bolinger, Dwight. 1985. Defining the Indefinable. I: Robert Ilson, red. *Dictionaries, Lexikography and Language Learning*: 69-73. Oxford.
- Collins COBUILD English Language Dictionary. 1987. London & Glasgow.
- Gellerstam, Martin. 1988. Verb Syntax in a Dictionary for Second-Language Learning. I: *Studies in Computer-Aided Lexicology*: 103-123. Stockholm.
- Hausmann, Franz Josef. 1985. Kollokationen im deutschen Wörterbuch: Ein Beitrag zur Theorie des lexikographischen Beispiels. I: Henning Bergenholtz & Joachim Mugdan, red. *Lexikographie und Grammatik: Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28.-30.6.1984*: 118-129. (Lexicographica: Series Maior, 3.) Tübingen.
- Kay, Martin. 1984. The Dictionary of the Future and the Future of the Dictionary. I: Antonio Zampolli & Amedeo Cappelli, red. *The Possibilities and Limits of the Computer in Producing and Publishing Dictionaries: Proceedings of the European Science Foundation Workshop, Pisa, 1981*: 161-174. (Linguistica computazionale, 3 (1983).) Pisa.
- LDOCE=Longman Dictionary of Contemporary English. 1978. London.
- Melčuk, Igor A. & Zholkovsky, Alexander K. 1984. *Explanatory combinatorial Dictionary of modern Russian. Semantico-Syntactic Studies of Russian Vocabulary*. /Åven rysk titel./ Wien.
- Nikula, Henrik. 1986. *Dependensgrammatik*. Malmö.
- Somers, Harold. 1987. *Valency and case in computational linguistics*. Edinburgh.
- Svensén, Bo. 1987. *Handbok i lexikografi*. Stockholm.
- Svensk ordbok*. 1986. Stockholm.
- Svenska ord med uttal och förklaringar*. 1984. Stockholm.
- Toporowska-Gronostaj, Maria. 1991. *Mot ett formaliserat valenslexikon*. (Opublicerad licentiatuppsats.) Göteborgs universitet: Institutionen för språkvetenskaplig databehandling.



Helge Niska og Jorunn Frøili

## Tolkordlistor på invandarspråk

Både i Sverige och Norge ger man ut särskilda terminologiska ordlistor avsedda främst för tolkar som arbetar inom samhällets tolkservice till invandrare. I denna uppsats behandlas principiella och praktiska överväganden vid urval av ämnesområden och termer samt principer för översättning av termer till olika språk. Vidare betonas vikten av feed-back från användarna samt behovet av nordiskt samarbete inom området terminologi på invandarspråk.

### 0. Inledning

Både Norge och Sverige har under de senaste decennierna haft en betydande invandring av människor från många länder med mycket varierande språklig och kulturell bakgrund. För att underlätta kontakten mellan invandrarna och värdsamhällets olika institutioner och organ har man inrättat tolkservice, bekostad av offentliga medel. I sina tolkade kontakter använder både myndigheter och invandrare fackuttryck och specialtermer, som tolken måste förstå och kunna översätta för att säkerställa en korrekt informationsöverföring. Dessa fackuttryck återfinns sällan i allmänspråkliga ordböcker - ja, faktum är att mellan nordiska språk och de flesta invandarspråken existerar det överhuvudtaget inga ordböcker.

### 0.1 Nyskapande arbete

Det har följaktligen funnits ett intresse hos både myndigheter, tolkar och tolkanvändare av att det skapas särskilda fackspråkliga ordlistor för tolkar. Vi har inga erfarenheter av denna typ av ordböcker sedan tidigare, så det är till stora delar frågan om nyskapande. En skillnad mot lexikografiskt arbete i allmänhet är att tolkordlistorna med nödvändighet, liksom andra terminologiska ordböcker, blir mer normerande än deskriptiva. Eftersom tolkordlistorna är avsedda att vara lätthanterliga arbetsverktyg för tolkarna, ställer de avsevärda krav på ekonomi vad gäller ordurval och grafiskt utförande. Och eftersom ordlistorna används vid tolkuppdrag inom sjukvård, rättsväsende och andra samhällseliga institutioner, ställs mycket höga krav på målspråkstermernas, dvs. översättningarnas, kvalitet.<sup>1</sup>

I Sverige har *Ordlista för tolkar* utkommit sedan 1970 och finns hittills i sjutton språkversioner. I Norge startade 1988 "Terminologiprojektet", som ska ta fram ordlistor för tolkar och översättare mellan norska och tolv invandarspråk.

Vi ska här redovisa några aspekter på arbetet med dessa ordlistor:

1. Urval av termer
2. Översättning av termer
3. Utformning av ordlistan
4. Feed-back och vidareutveckling

## 1. Urval av termer

### 1.1 Avgränsning av ämnen

Syftet med ordlistorna skall vara att tillgodose behovet av terminologi i invandranas kontakter med värdsamhället. Det har främst varit samhällets behov av korrekt och effektiv information till invandrarna som har styrt arbetet. Ordlistorna tar således upp endast svenska respektive norska begrepp, inte specifika begrepp från tolkspråksländerna.

Första steget har varit att utarbeta en basordlista, som sedan översätts till de aktuella invandrarpråken. Man har begränsat sig till de huvudområden där de flesta tolkuppdragen utförs. Följande ämnesområden är således inkluderade i den norska listan (den svenska har något färre områden): offentlig förvaltning, invandrarpolitik, utbildning, arbetsliv, skatt, finans (post och bank), bostad, trafik (körkortspröv), socialtjänst och försäkring, sjukvård inkl. psykiatri, juridik.

### 1.1 Excerpering av termer

En principiellt viktig fråga är hur man avgör vad som behöver tas med. Det kan tänkas att också ord från allmänspråket upplevs som besvärliga för en invandrare med annan kulturell bakgrund. Även individuella kunskapsskillnader spelar in. Är "barnehage", "øye" och "choke" så kända och självklara att de bör slopas? Excerperingsarbetet har inte oväntat varit svårt och ständigt kritiserat.

Tre huvudintressenter har påverkat ordlistornas ordförråd:

- tolkarna själva, som före tolkordlistornas tillkomst såg bristerna i eventuellt existerande allmänspråkliga ordböcker och själva tillverkade specialordlistor för sina egna behov,
- tolkanvändarna, d.v.s. i första hand fackmän inom olika områden där tolkning ägde rum,
- samhället, som *dels* behövde få ut egen information även till invånare som inte behärskade landets språk, *dels* var bundet av lagstiftning om rätt till tolkning hos myndigheter.

Som källor har använts offentliga broschyrer, *Samfunnsboka*, *Samhällsguiden* och liknande, undervisningsmaterial från tolkkurser, såsom kompendier och läroböcker, samt terminologilistor. I Norge har Rådet for teknisk terminologi (RTT) gjort excerperingen.

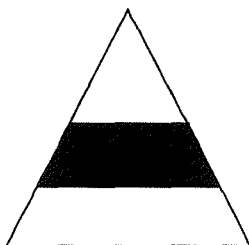
Endast i begränsad utsträckning har det varit möjligt att använda redan befintliga ordböcker, eftersom denna typ av lexikon inte funnits tidigare.

### 1.2 Avgränsning av ordförrådet

#### 1.2.1 Hanterbarhet

En första utgångspunkt har varit, att ordlistan ska vara hanterlig och lätt att ta med sig på tolkuppdrag. Den ska inte belastas med ord och termer som kan anses tillhöra allmänspråket eller höra till allmänbildningen. Det finns också en gräns åt andra hållet - alltför specifika termer kan inte tas med av praktiska skäl. Det är självklart också en kostnadsfråga. Den svenska listan innehåller ca 6 500 uppslagsord, och den norska kommer att ha ca 4 000 definierade termposter plus ett antal hänvisningsposter.

Avgränsningsprinciperna kan illustreras som i fig. 1. Om triangeln representerar det sammanlagda ordförrådet, där basen är allmänspråkets ord och triangelns topp är fackspråkets termer, ligger tolkordlistans ordförråd som ett band någonstans ovanför basen och upp till ett område innan triangeln blir alltför smal, där ordförrådet alltså blir alltför specialiserat.



**Figur 1** Fig. Avgränsning av ordförrådet i *Ordlista för tolkar*

### 1.2.2 Sammansättningar

För att spara plats har man i *Ordlista för tolkar* ofta låtit bli att ta med huvudordet, men väl tagit med sammansättningar där ordet förekommer. Följande uppställning under bokstaven T visar principen och illustrerar också vilka sammansättningar som ansetts viktiga att ta upp. I princip försöker man undvika alltför enkla, "uppenbara" sammansättningar, men gränserna är svåra att dra. Ordet "tand" finns alltså inte med, men däremot:

tandben  
 tandbrygga  
 tandhals  
 tandhygienist  
 tandkött  
 tandköttsficka  
 tandlossning  
 tandprotes  
 tandreglering  
 tandröta  
 tandsten  
 tandställning  
 tandtekniker  
 tandvårdsersättning  
 tandvårdsförsäkring

Efter "tantiem", "tapp (opt)" och "tariffjön" kommer så nästa sammansättningsgrupp:

tarminklämning  
 tarmkäv  
 tarmludd  
 tarmvred

Man kan diskutera lämpligheten av att inte ta med huvudordet som uppslagsord. Under alla förhållanden strider det mot normala principer vid ordboksframställning. Någon kritik av detta tillvägagångssätt har dock inte framkommit från brukarna.

### 1.2.3 Ordgrupper, fraser

I *Ordlista för tolkar* har strävan varit att få med ordgrupper och fraser som kan betecknas som fasta eller etablerade inom t. ex. förvaltningen eller något fackområde. Således finns "föräldrapenning för tillfällig vård av barn" med som uppslagsord, men inte "vård av barn". Den första frasen tillhör den familjerättsliga terminologin, medan den andra är ett vardagligt begrepp, som kan uttryckas på många andra sätt, t.ex. "ta hand om barn", "skötsel av barn" - det är alltså ingen egentlig term. Av samma skäl finns såväl "vårdnad om barn" som "gemensam vårdnad" med, men inte "vårdnad".<sup>2</sup>

### 1.2.4 Endast dagens terminologi

Tolkarna arbetar i högsta grad i nuet och måste alltså alltid ha tillgång till senaste nytt i fråga om termer på olika områden. Vid revidering och nytillverkning av *Ordlista för tolkar* låter man föråldrade termer finnas kvar endast om termen ifråga fortfarande är i allmänt bruk. Man har t.ex. mönstrat ut ordet "kräfta" och har bara kvar beteckningen "cancer". Den föråldrade domstolstermen "lyckta dörrar" (används väl numera endast i tidningsartiklar och kriminalromaner) har också fått stryka på foten, och där har man inte haft något behov att sätta in den moderna benämningen "låsta dörrar" istället.

### 1.2.5 Namn på myndigheter och organisationer; yrkesbeteckningar

Policyn har varit att ta med så få benämningar på myndigheter och organisationer som möjligt - administrativa organ, departement och politiska partier byter namn eller försvinner, så ock deras förkortningar.

Yrkesbeteckningar förekommer i den mån de representerar en samhällsfunktion eller institution, t. ex.

barmorska  
hemsamarit  
hemspråkslärare  
socialsekreterare  
tandhygienist  
åklagare

### 1.2.6 Kulturellt/språkligt betingad anpassning

Grundprincipen har varit att alla termer i basordlistan skall översättas. Men alla termer har inte känts lika angelägna att ta med i alla språkversioner. När den finska versionen av *Ordlista för tolkar* reviderades 1985, ville översättaren utöka basordlistan med namnen på ett antal allergiframkallande växter i den nordiska floran, eftersom allergier av bl.a. denna sort

var mycket vanliga. De som var ansvariga för en samtidig revidering av den grekiska listan protesterade med hänvisning till att dessa växter ofta saknade benämning på grekiska och att greker i allmänhet ändå inte kunde namn på växter!

Ett annat avvägningsproblem gäller det som brukar kallas "samhällstermer", dvs. termer som beskriver företeelser inom samhällsservice och samhällsadministration. För invandrare från länder med mycket annorlunda samhällssystem än vårt, kan även begrepp som för oss är självklara ställa till bekymmer. Man bör då kunna vara flexibel. I den finska tolkordlistan finns t ex "sjuksköterska" och "sjukstuga" inte med - det gör de däremot i den nyutgivna persiska tolkordlistan.

## 2. Översättning av termer

Härnedan beskrivs sex strategier för översättning av termer. Vid översättning av basordlistorna till olika språk, måste man givetvis använda de strategier som passar det språk man översätter till. Man är bunden av de konventioner och traditioner, inte minst av språkpolitisk art, som finns hos respektive språksamfund vad gäller termbildning.

1. *Direkt motsvarighet finns på det andra språket.* Om det finns en termekvivalent, dvs. en term som på det andra språket står för exakt samma begrepp som utgångsspråkets term, använder man den. Detta är vanligt när det gäller t.ex. internationella standarder och nomenklaturer, men ovanligt i fråga om samhällstermer. Även för två så kulturellt näraliggande samhällen som det svenska och det finländska har det uppstått ett stort antal finska termer som enbart används i Sverige för att beskriva företeelser som saknas i Finland.<sup>3</sup>

Om exakt motsvarighet saknas, kan man ibland ta till en "ungefärlig" ekvivalent (ibland kallad "översättningsekvivalent"), dvs. en term som står för ett likartat eller näraliggande begrepp.<sup>4</sup> Ex.: le baccalauréat - sv. "studentexamen"-en. "'A' level"<sup>5</sup>

2. *Översättningslån.* Om språkets struktur eller konventioner tillåter det, översätts sammansatta ord eller ordgrupper direkt, komponent för komponent. Detta är en vanlig strategi bl. a. vid översättning av kulturspecifika termer såsom samhällstermer, typ la Communauté Européenne - the European Community - Europeiska Gemenskapen - Det Europeiske Fellesskap.

I *Ordlista för tolkar* har folkhögskola på engelska blivit "folk high school". För ett engelskt eller amerikanskt öra torde denna term inte säga speciellt mycket. Frågan är om inte denna direktöversättning borde kompletteras med något begrepp som är känt i målspråkskulturen, t.ex. "adult education centre", alltså en ungefärlig ekvivalent enligt strategi. En sådan komplettering kan man kalla översättningsdublett, se punkt 6.

3. *Översättning av förklaringen av källtermen.* Detta är vanligt i de äldre versionerna av *Ordlista för tolkar*, men undviks i de nyare versionerna till förmån för strategi 2 eller 5. Även om förklaringen görs så komprimerad som möjligt, blir kommunikationen ändå tungrodd. Dessutom får översättningen i och med att den är en förklaring en annan precisionsnivå än källtermen. Ex.: le baccalauréat - en. "the French secondary school leaving examination".

4. *Termen överförs direkt (direktlån).* Direktlån är tämligen vanliga i invandrades talspråk. Typexempel: no. grunnfag - ar. "grunnfag", sv. försäkringskassa - pers. "försäkringskassa". Ofta anses dylika lån vara uttryck för slarvigt och ovårdat språkbruk. Men även i skriftspråk och vårdat språk är direktlån vanliga. De kan förekomma som rena *citatlån*, såsom när de

engelska orden input, output, engineering, know-how tas in "som de står" i våra nordiska språk. Vanligare är dock *modifierade direktlån* där det inlånade ordet anpassas till målspråkets fonetik och morfologi. Ex.: medicinsk och teknisk terminologi: afasi - fi. "afasia", scintografi - fi. "skintografia".

5. *Nybildning*. Nybildningar av termer innebär inte nödvändigtvis att man tillverkar nya ord, man kan också konstruera nya sammansättningar och ordgrupper av redan existerande ord. Det kan också innebära att man låter ett t.ex. föråldrat eller dialektalt ord få en ny betydelse. Nybildning av termer är en helt naturlig process inom alla språk, som sker på olika sätt beroende på språkens struktur och språkpolitiska traditioner.

Ett modernt, av många illa omtyckt, exempel på nybildning inom samhällsterminologin i Sverige är det engelska approximation som i den svenska översättningen av EG:s samlade regelverk har fått benämningen "tillnärmning".

6. *Översättningsdubletter*.<sup>6</sup> Översättningsdubletter innebär en kombination av två av ovanstående strategier. Ex. le baccalauréat - sv. "baccalauréat, fransk studentexamen" (kombination av citatlån och förklaring), reaktion - fi. "reaktio, vastavaikutus" (modifierat direktlån och översättningslån). Till och med *översättningstripletter* kan förekomma, såsom i följande exempel ur *Ordlista för tolkar*: EKG, elektrokardiogram - fi. "EKG, elektrokardiogrammi, sydän(sähkö)käyrä, 'sydänfilmi'". Först kommer ett direktlån, modifierat för att passa finsk morfologi, därefter ett översättningslån ("hjärt(elektricitets)kurva") och sist inom citationstecken lekmannauttrycket "hjärtfilm".

### 3. Utformning av ordlistan

Vilka data ska tas med i ordlistorna? Detta är en avvägningsfråga mellan hanterbarhet (ordlistan ska vara möjlig att ha med vid tolkningsuppdrag och det ska vara lätt att snabbt slå upp en term mitt under en tolkning) och brukarnas behov av definitioner, synonymer, grammatisk information m.m. om termerna.

I det norska terminologiprojektet använder man sig av den mall som Rådet for teknisk terminologi (RTT) använder sig av. Termposterna är uppställda i en *databas* i alfabetisk ordning. Det finns två slag av poster, *definierade termposter* och *hänvisnings-/synonymposter*. De sistnämnda innehåller bara en hänvisning till en definierad termpost.

För de definierade termposterna ges följande information:

1. Uppslagsordet (själva termen)
2. Översättning av uppslagsordet (skrivs in på samma rad som uppslagsordet<sup>7</sup>)
3. Grammatisk upplysning (ordklass, ev. genus)
4. Eventuella synonymer
5. Definition av uppslagsordet
6. Anmärkningar (tilläggsinformation, hänvisning till annan termpost)
7. Ämnesområde (kallas *grupp*, även undergrupper anges)
8. Bokstavskod för identifiering (finns endast i databasen, skrivs ej ut i de tryckta ordlistorna)

Tolkordlistorna förses med två bilagor: ett *ämnesregister* och ett *målspråksregister*. I ämnesregistret står de tidigare nämnda tolv ämnesområdena uppställda var för sig med tillhörande uppslagsord (enbart uppslagsorden). Målspråksregistret är ingen vändning av ordlistan, utan ett försök att där det är språkligt möjligt ge en ledtråd till den norska termen. För full information om uppslagsordet måste brukarna i båda fallen gå till den alfabetiska huvudordlistan.

#### 4. Feed-back och vidareutveckling

##### 4.1. Kritik från användarna

Det ligger i sakens natur, att den typ av ordlistor som vi beskrivit fungerar normativt. Även om intentionen skulle vara att försöka beskriva mötet eller skärningspunkten mellan två språk, är det klart att den eller de översättningar som valts fungerar som rättesnören i många sammanhang: i ordinarie användning på tolkuppdrag, vid skriftlig översättning, i tolk- och översättarutbildning och även vid auktorisationsprov för tolkar och översättare.

Mot denna bakgrund är det viktigt att släppa fram kritik och kommentarer till tolkordlistorna. Ordlistorna måste ständigt vara föremål för bearbetning och revidering. De norska listorna tas fram i samarbete mellan Utlendingsdirektoratet, Rådet for teknisk terminologi (RTT) och Avdeling for tolkeutdanning vid universitet i Oslo. Tillsammans representerar dessa organ avsevärda kunskaper och erfarenheter av terminologiskt arbete, invandrarservice och tolkutbildning, men en remissrunda är ändå viktig innan de färdiga ordlistorna ges ut. På det sättet kan man på ett tidigt stadium beakta synpunkter på uppslagsord, översättningar m.m.

De svenska tolkordlistorna har under den tjugoårsperiod de hittills getts ut genomgått en del förändringar och uppdateringar. Mycket av de förändringar som skett har haft sin grund i användarnas påpekanden. I början av åttiotalet lät Kommerskollegium, som är den myndighet som auktoriserar tolkar och översättare, granska de flesta tolkordlistorna på grund av att häftig kritik riktats mot dem. I de flesta fall visade det sig att kritiken varit överdriven, men i ett par fall (finska och grekiska) ledde det till att ordlistan skyndsamt reviderades.

##### 4.2 Utvecklingsarbete

På de språk där tvåspråkiga ordböcker inte funnits alls tidigare, har det visat sig att arbetet med tolkordlistorna har gett erfarenheter och kunskaper som varit värdefulla när man så småningom börjat framställa allmänspråkliga ordböcker. Inte minst har man fått fram ett antal översättare som med lämplig vägledning och lexikografisk skolning kan vara utomordentliga resurser för vidare lexikografiskt arbete på främmande språk.

Ett problem vid utarbetande av ordlistor på invandrarspråk är att det i de allra flesta språk saknas en självklar auktoritet, som skulle kunna avgöra vad som är "rätt eller fel" i en översättning av en term eller ge råd och tips i terminologiska frågor till dem som översätter listorna. Inom Norden har vi på invandrarspråken brist på kunniga terminologer, terminologiska utvecklingsorgan eller språkvårdande

institutioner i allmänhet. Vi saknar också universitetsutbildning och forskning i många språk.

En utökad internationalisering skulle lösa en del av problemen. På många språk sker ett omfattande terminologiskt arbete i andra länder och inom internationella organ, men vi har för lite kontakt med detta arbete. Det gäller till exempel arabiska, persiska, kurdiska, indiska språk, malayiska och swahili. Det vore till att börja med en angelägen uppgift att få igång ett samnordiskt terminologiskt utvecklingsarbete på våra invandrarspråk. Det skulle även vara till gagn för de inhemska språken och skulle avsevärt höja den terminologiska och lexikografiska kompetensen i våra länder.

### Noter

1. "Översättning" kallar vi för enkelhets skull den term som står på målspråket i ordlistan, oavsett hur målspråkstermen skapats (se avsnittet om översättning av termer).
2. *Ordlista för tolkar* började ges ut 1970, långt innan fackspråksforskning och lexikografi blev etablerade vetenskaper. Inte förrän på senare år har kollokationer och fackspråkliga fraser uppmärksammats inom terminologiforskningen. Kjær (1990) talar om tre typer av fackspråkliga fraser:
  - free or predictable word combinations
  - semi-fixed or idiosyncratic word combinations
  - fixed or phraseological word combinations
 I tolkordlistan har man uppenbarligen strävat att i första hand ta med fraser av de två sistnämnda typerna.
3. Se Kieliviesti (1985) för en redogörelse för sverigefinskt terminologiarbete.
4. Newmark (1988) kallar detta "cultural equivalents". Williams (1990) redogör mera välstrukturerat för Newmarks typologi.
5. Följande förkortningar för språk används i detta avsnitt: ar. arabiska, en. engelska, fi. finska, fr. franska, no. norska, pers. persiska, sv. svenska.
6. "Translation couplets" enligt Newmark (1988). Se även Williams (1990).
7. Tekniska problem, som dock inte är oöverbärliga, uppstår vid användning av främmande, särskilt icke-latinska alfabet såsom det arabiska för persiska och urdu.

### Litteratur

- Kieliviesti nr. 3-4 1985. Stockholm: Sverigefinska språknämnden.
- Kjær, Anne Lise. 1990. "Phraseological research - State-of-the-art. Methods of describing word combinations in language for specific purposes." I: *Terminology Science & Research. Journal of the International Institute for Terminology Research (IITF Journal)*. Vol. 1, no. 1 - 2, 3-20.
- Newmark, Peter. 1988. *A Textbook of Translation*. New York etc: Prentice Hall.
- Ordlista för tolkar*. 1979-. Stockholm: Liber Utbildningsförlaget.



Williams, Jennifer. 1990. "The Translation of Culture-Specific Terms". I: *Lebende Sprachen*. Nr. 2/90, 55-58.



Del 5



Lars Svensson

## Om SAOB:s material och en bibliografisk databas

SAOB:s excerptsamling omfattar ca 7,5 miljoner språkprov. Med excerperingen, som är selektiv, har man eftersträvat kronologisk spridning av materialet och samtidigt velat täcka så mycket som möjligt av svenskt språk (olika stilsfärer och olika språk i språket). Excerptsamlingen, som omfattar ca 20.000 källor, innehåller inte bara tryckt material utan också ett omfattande handskriftsmaterial. Stickkontroller i bibliografiska arbeten visar att SAOB väl täcker svenskt boktryck från 1500-talet till våra dagar. — Sedan några år tillbaka pågår vid SAOB-redaktionen ett projekt med syfte att skriva och optiskt läsa in SAOB:s källförteckningar i en bibliografisk databas med en mängd sök- och sorteringsmöjligheter. Bibliografin kommer att kunna länkas till den databas som Språkdata i Göteborg håller på att upprätta över de tryckta banden av SAOB.

*Svenska Akademiens ordbok* (SAOB)<sup>1</sup> är en historisk-nuspråklig ordbok som behandlar svenskt skriftspråk från 1521 till nutiden. Hittills (1991) är 30 volymer utgivna. Band 1 publicerades 1898, band 30 1989. Jag skall nedan mycket översiktligt försöka ge en bild av SAOB:s material. Mina huvudkällor härvidlag har varit kapitel 1 och 2 i Ekbo-Loman, *Vägledning till Svenska Akademiens ordbok*, skrivna av före ordbokschefen Sven Ekbo, och uppsatsen "Om SAOB:s material" (i *Nysvenska Studier* 1983), skriven av Sture Hast, som efter 30 års tjänstgöring som ordboksredaktör vid SAOB nyligen gått i pension. Jag kommer också att kort presentera SAOB:s *Källförteckning. Supplement II* (1990), som — i motsats till SAOB:s källförteckningar från 1939 och 1975 — är en bibliografi, inte bara en alfabetisk källförteckning.

### 7,5 miljoner egna excerpter

SAOB:s egen excerptsamling omfattar i dag ca 7,5 miljoner språkprov. Excerpterna, som är försedda med uppgift om källans namn (författare, boktitel o.d.), sida eller paragraf, datering o.d., förvaras i alfabetisk ordning efter de olika uppslagsorden i kapslar av papp. Varje excerptlapp omfattar ett enda språkprov. Excerpter som avser samma ord ordnas så att det äldsta språkprovet ligger först, det yngsta sist. Språkproven är skrivna för hand (det stora flertalet) eller på skrivmaskin eller utgörs av en fotograferad boksida o.d. eller urklipp av tidningstext som klistrats på excerptlapparna.

### Material från andra samlingar

SAOB får också visst material från andra samlingar: språkforskaren Esaias Tegnér's språkprovssamling (flera tusen excerpter), arkivmannen och urkundsutgivaren Victor Granlunds ordspråkssamling, språkforskaren Olof Östergrens språkprovssamling (ca 250.000 excerpter) till *Nusvensk ordbok*, 10 band med riklig fraseologi, Tekniska Nomenklaturcentralen (TNC), ordförteckningar med hänvisningar till källor, Svenska Språknämnden (ca 1.000 excerpter per år), — Språknämnden får i gengäld visst, nytillkommet material från

SAOB —, Rolf Dunås klipp- och excerptsamling (flera tusen excerpter). — Dunås var lärare och svarade under flera år på språkfrågor i en dagstidning. Redaktionen får också sporadiska bidrag av enskilda forskare och ordälskare.

Här skall också nämnas en serie konkordanser på mikrofiche utarbetade av Språkdata vid Göteborgs universitet: press 76, omfattande 1, 3 miljoner löpande ord, romaner 76-77, 80-81, tillsammans omfattande 9,3 miljoner löpande ord ur 129 romaner, Svensk författningssamling 77-80, ca 1/2 miljon löpande ord. Språkdata-konkordanserna är ett viktigt, modernt material. De har icke excerperats men används för anskaffande av moderna belägg.

Räknar vi samman SAOB:s egen språkprovssamling och materialet från andra institutioner och samlingar får vi en imponerande och solid bas för redigeringen av ordboken. Följande siffror kan ge en antydning om materialets stora omfattning. Band 30 (*spår-stockna*) upptar 1.615 tättskrivna spalter, vilket motsvarar knappt 8 sidor i SAOL 1986, dvs. ca 202 spalter (eller ca 16.000 excerpter) i SAOB motsvarar ungefär 1 sida i SAOL. Det största parti som — hittills — delats ut är partiet "stor", dvs. substantivet och adjektivet "stor" jämte sammansättningar och avledningar. Det omfattade över 18.000 excerpter. Det tog ca 18 månader för redaktören att skriva första manus. — SAOB:s excerptmaterial står sig gott vid en jämförelse med utländska ordböcker av motsvarande typ: Ordbog over det danske sprog, ca 2,5 miljoner excerpter, Oxford-ordboken (*New English Dictionary*), ca 6 miljoner, *Thesaurus Linguae Latinae*, ca 10 miljoner, *Orðabók Háskólans*, ca 2,5 miljoner.

### Säkerhetsfilmning av språkprovssamlingarna

Av säkerhetsskäl har den del av materialet som ännu icke redovisats i tryck (*ymb-ö*) mikrofilmats med ett datorsystem som medelst en mikrofilmläsare möjliggör återsökning av språkproven (ca 1,5 miljoner). En kopia av mikrofilmerna går att anskaffa från SAOB. I första hand borde universitetsbibliotek och språkliga institutioner vara intresserade av att ha tillgång till det utryckta materialet. Observera att blott en ringa del (ca 10-15%) av det väldiga excerptmaterialet redovisas i tryck (för att ordbokens omfång skall begränsas). Så småningom kommer även materialet till den tryckta delen av ordboken att mikrofilmats.

### Excerptering utanför och inom redaktionen

Excerpteringen — jag förbigår här ordbokens äldre historia — påbörjades i mitten av 1880-talet och pågår alltjämt. Den utförs i dag av 6-7 fasta excerptister utanför redaktionen, som excerpterar tidningar, tidskrifter, årsskrifter, skönlitteratur etc. Ordbokschefen och chefkontrollanten väljer ut den litteratur som skall excerpteras, ofta på förslag av redaktör, kontrollant eller annan medarbetare. Stundom föreslår också någon av SAOB:s facksakkunniga böcker för excerptering. Ofta får någon erfaren redaktör eller chefkontrollanten i uppdrag av ordbokschefen att göra en grov förexcerptering för att utvärdera om skriften i fråga kan tillföra något nytt utöver det som redan finns i språkprovssamlingarna. En *Hushållsbok*, t.ex., från slutet av 1700-talet kanske ratas som excerpteringstext. Belägg på ord som *öl*, *humle*, *salt*, *socker*, *gryn*, *ärter* osv. finns nämligen — med största sannolikhet — i redan excerpterade hushållsböcker från 1500- och 1600-talen.

Materialet kompletteras emellertid också av redaktionen, framför allt av assistenterna i den förberedande redigeringen, som genomgår bortemot 250 verk (ordböcker, ordlistor,

konkordanser, uppslagsböcker, florum, apotekartaxor, riksdagsprotokoll m.m.), men även av de assistenter som har till uppgift att kontrollera redaktörernas artiklar (beläggens datering, betydelseinnehåll och stavning, definitionernas hållbarhet i förhållande till språkproven osv.). Även redaktörerna själva kompletterar i viss utsträckning materialet under arbetets gång, ofta i samband med specialstudier av ett visst fackområde.

### Excerptisterna

Under perioden 1894—1924 hade excerptisterna en egen chef, språkmannen Theodor Hjelmqvist. Nya excerptister fick under denna tid också en (alltför) utförlig och detaljerad instruktionsskrift på 13 sidor som hjälpmedel, "Anvisningar till insamlande af språkprof för Svenska Akademiens ordboksarbete". I dag får en ny excerptist instruktioner av chefkontrollanten, muntligen eller medelst brev. Viktig information om SAOB:s arbete och material meddelas också på de seminarier med föredrag och diskussioner, som arrangeras av SAOB-redaktionen med jämna mellanrum och till vilka excerptister och facksakkunniga särskilt inbjuds. — Sammantaget har SAOB hittills anlitat flera hundra, ofta handplockade, excerptister med fin språkkänsla och språkuppfattning, till yrket ofta skolmän och lärare, bibliotekarier, präster eller vetenskapsmän inom skilda fack och fält. Även pensionerade SAOB-medarbetare har anlitats och anlitas ännu som excerptister.

### Noggrann kontroll av språkproven

Alla språkprov som medtages i trycket blir föremål för omsorgsfull kontroll. Excerpterna är i allmänhet tillförlitliga vad beträffar själva texten. Inte så sällan förekommer det dock korrupta textställen, utgivningsfel, skriv-, läs- och tryckfel, vilka kan vålla redaktör och kontrollant stort huvudbry. Felaktiga käll- och sidhänvisningar rättas till i manus av kontrollanterna, som också — om möjligt — flyttar belägg som icke excerperats i originalupplaga till en sådan. Ibland har excerptisten varit för snål med kontext, så att betydelsen av textstället inte klart framgår. I sådana fall måste man läsa en längre passage i källan och då är man tacksam för att redaktionen är inrymd i Lunds universitetsbibliotek, där redaktionspersonalen har tillstånd att röra sig fritt i magasinerna. — Före 1924 granskades excerptistens arbete mycket noga av excerptistchefen. Särskilt ordurvalet blev föremål för kritik eller beröm. Inlämnat excerptmaterial granskas också i dag men allvarlig kritik mot excerptisten är sällsynt.

### Selektiv excerptering

Det ordförråd som behandlas i SAOB är det svenska skriftspråkets från 1521 fram till vår egen tid. I källorna som excerperas väljer excerptisten ut 3-5 ord per sida. Vissa källor har emellertid blivit — eller blir — föremål för en betydligt utförligare excerptering (t.ex. Bibeln, Olaus Petris skrifter, Linnés skrifter), dock aldrig så, att alla löpande ord i en text blivit excerperade. Det finns naturligtvis en risk med selektiv excerptering: de mest frekventa orden och uttrycken kan bli dåligt representerade. Excerptisten går lätt förbi de vanliga orden och noterar i första hand språkliga "egendomligheter". Å andra sidan är det betydligt lättare för redigerande redaktör och kontrollant att komplettera bristen på "vanliga" ord och uttryck än

på "ovanliga", om så skulle behövas. SAOB iakttar och måste iaktta restriktivitet i fråga om utökande av materialet. Ett alltför stort material hindrar starkt utgivningstakten. Ett franskt ordboks företag, som med dator hade anskaffat 90 miljoner excerpter, använde i realiteten bara 6 miljoner vid redigeringen. SAOB:s språkprovssamling utökas sedan ca 30 år tillbaka årligen med ca 10.000-12.000 excerpter. Excerpteringen sker sedan 1959 likformigt över hela alfabetet. En del av det nytillkomna materialet kommer således att behandlas i ett kommande *Supplement* (eller i en ny upplaga av ordboken).

### Urvalsprinciper

I SAOB behandlas svenskt skriftspråk från nästan 5 sekel men icke utan inskränkning. Gränserna för vad som kommer med eller utesluts blir ibland flytande. Följande kategorier av ord medtas i regel icke i SAOB:

- a) namn utan användning såsom appellativ, dvs. personnamn och ortnamn,
- b) tillfälliga ord (särskilt sammansättningar) och bildningar,
- c) extremt fackspråk, särskilt sådant som ej nått utanför den trängre fackkretsen,
- d) dialektord i dialektordböcker eller dialektundersökningar eller i text som är skriven helt på dialekt,
- e) främmande ord som ej har svensk böjning och ej vunnit hävd i svenskan,
- f) rena citatord från främmande språk.

### Kommentar

Viktiga litteraturverk och kända litteraturställen prioriteras. Man är i allmänhet också generösare mot belägg från 1500- och 1600-talen än mot senare belägg. Kryptogamnamm belagda endast i 1900-talsfloror förbigås liksom i regel termer som utgör översättning från främmande språk och endast är belagda i referat av utländska undersökningar eller förhållanden i källor som *ÅrsbVeta*, *KrigVAH* och *KrigVAT*. Individuella ordskapelser (t.ex. hos Olof Högberg eller Carl August Hagberg, i hans Shakespeare-översättningar) förbigås i allmänhet. Sammansättningar till sammansättningar utgallras strängt. Använder en författare dialektord i berättande framställning tas de i regel med.

### Källorna

Medeltidens svenska skriftspråk är täckt genom Söderwalls och Schlyters ordböcker. I svensk språkhistoria brukar man räkna den äldre nysvenskan från 1526, *Nya Testamentets* utgivningsår, men startåret för SAOB:s excerptering har satts till 1521, för att *GIR* skall komma med i sin helhet — de äldsta akterna där är daterade 1521. SAOB:s källförteckningar — 1898 (i första bandet), 1939, 1975, 1990 (den sista är ännu icke publicerad) — upptar ca 20.000 titlar men antalet texter är betydligt större. En källa som omfattar flera band eller volymer uppföres nämligen i *Källförteckningarna* som en. Så omfattar t.ex. källan *GIR* 29 band, *PH* 15 band, *OxBr.* 12 band. Likaså anges en tidning eller tidskrift eller årsbok o.d. som en enda källa, även om flera årgångar har excerpterats. Omvänt förekommer naturligtvis också att åtskilliga texter, t.ex. almanackorna under 1600-talet, bara består av ett 10-tal sidor. I



*Källförteckningen* 1939 har arbeten, som endast citeras någon enstaka gång, ej medtagits. I *Supplement I* och *II* har däremot alla citerade källor medtagits, även sådana, varifrån endast enstaka belägg hämtats. Sture Hast (*Nysvenska Studier* 1983, s.173 f.) har efter en stickprovskontroll kommit fram till att endast 1/5 av de citerade texterna i *Supplement I* verkligen är excerperade, dvs. att excerptisten har skrivit av 3-5 belägg per sida. I flertalet källor i *Supplement I* har alltså blott enstaka belägg excerperats.

Med excerpteringen har man — samtidigt som man eftersträvat kronologisk spridning av materialet — velat täcka så mycket som möjligt av svenskt språk, t.ex. olika stilsfärer och olika språk i språket (bygdemålsfärgat språk, vardagligt språk, högtidligt språk, religiöst språk, ålderdomligt eller arkaiserande språk, slang, *arkeologi, botanik, ekonomi* etc.). *Källförteckningarna* omfattar ca 150 genrer och ämnesområden: vitterhet, teologi, språkvetenskap, förordningar, kanslihandlingar, astronomi, kemi, militärväsen, sport, brev, dagböcker, domböcker, tänkeböcker (dvs. stadsböcker innehållande anteckningar av judiciell, kommunal och kameral art), biografi, botanik, zoologi, reseskildringar, tidningar, tidskrifter. Osv. Även barn- och ungdomsböcker, t.o.m. serietidningar, är representerade. Excerptsamlingen innehåller inte bara tryckt material utan också ett mycket omfattande handskriftsmaterial (räkenskapshandlingar, bouppteckningar o.d.). Detta är ett mycket viktigt material. Kulturord, vanligen lånord, uppträder ofta först i tullhandlingar och slottsräkenskaper och får sedan vidare spridning. I bouppteckningarna anträffas äldre benämningar på redskap, verktyg, husgeråd o.d.

I detta sammanhang kan nämnas att Bibliotekstjänst i Lund, dit praktiskt taget allt svenskt tryck skickas för recension, kontinuerligt (2 gånger i månaden) sänder oss sambindningsrecensionerna (som utgör underlag för bibliotekens inköp av litteratur), varigenom vi fortlöpande underrättas om allt nytryck i Sverige.

Fig. 1 Några viktigare, utförligt excerperade källor.

1. Bibeln och psalmboken.
2. GIR (Konung Gustaf den förstes registratur)
4. PH (Utdrag utur alle ifrån .. 1718 utkomne publique handlingar)
5. Tänkeböcker
6. Encyklopedier
7. Protokoll från ständsriksdagarna
8. Axel Oxenstiernas skrifter och brevvetling
9. Svenska Akademiens handlingar
10. Vetenskapsakademiens handlingar
11. Tidningar och tidskrifter
12. Handskrifter.

(Efter Sture Hast, *Nysvenska Studier*, 1983, s.180)

### Källornas fördelning på genrer

Sture Hast har i *Nysvenska Studier* 1983, s.174 f. gett en översikt av källornas fördelning på olika genrer eller fackområden. Han har granskat *Källförteckningen* i sin helhet från 1898 (24 sidor i band 1), 50 spalter i *Källförteckningen* tryckt 1939 och *Källförteckning. Supplement I* tryckt 1975. I tabellen nedan har jag lagt till siffror ur *Källförteckning. Supplement II* från 1990. Jag har också ändrat uppställningen. Den största genren (i *Supplement II*) kommer först.

Fig. 2 Jämförelse mellan källförteckningarna med avseende på genrefördelning (i procent).

Genrer	1898	1939	1975	1990	antal excerp. texter 1990
vitterhet	29,0	21,8	42,0	18,1	329
handskrifter	0,1	1,2	1,7	7,5	136
tidskrifter	2,3	1,8	1,1	6,6	121
teknik	0,6	1,4	2,5	6,3	114
juridik	1,6	2,1	0,7	4,8	88
språkvetenskap	4,0	3,8	2,4	5,2	95
biografi	2,7	8,3	4,2	4,3	78
ekonomi	2,5	2,5	4,3		79
sport	1,6	0,8	2,0	3,9	71
historia	6,0	5,5	5,1	3,7	68
naturvetenskap	1,5	1,1	1,2	2,9	53
officiellt tryck	1,7	3,3	0,5	2,7	54
teologi	12,5	6,5	3,1	2,6	47
tidningar	1,6	1,3	0,7	2,2	40
geografi	1,4	3,0	1,8	2,2	40
medicin	2,2	3,3	2,0	2,1	39
matematik, fysik, kemi	1,1	3,4	1,3	1,6	29
litteraturvetenskap	1,1	2,2	2,5	1,5	27
ordböcker	4,9	1,1	2,0	1,2	21
botanik och trädgård	1,2	3,2	0,8	1,2	23
resor	1,6	2,1	3,2	0,9	17
militärväsen	4,1	2,1	2,3	0,6	11
sjöväsen	1,0	0,3	0,4	0,5	9
åkerbruk	1,0	2,4	1,3	0,3	5
pedagogik	1,4	2,4	0,8	0,2	3

### Kommentar

Som framgår är *vitterhet* fortfarande den största genren. Att siffrorna är relativt små 1898 och 1939 beror säkerligen på att åtskilliga "samlade verk" tagits upp som en enda källa. *Handskrifterna* i *Supplement II* (1990) är påfallande många. Detta torde bero på att samtliga källskrifter upptas här, även sådana som endast excerperats för något enstaka belägg eller endast citerats någon enstaka gång. I *Supplement II* har siffrorna för *tidskrifter*, *teknik*, *juridik*, *språkvetenskap* (i mina siffror ingår även språkvetenskapliga *ordböcker*), *biografi*, *ekonomi*, *sport*, *naturvetenskap*, *officiellt tryck* och *tidningar* ökat. Siffrorna för *teologi*, *militärväsen* och *pedagogik* fortsätter att rasa. Även *reseskildringarna* har avtagit.

### Hur täcker materialet svensk bokproduktion?

#### 1500- och 1600-talen

En snabb genomgång i Collijns 1500- och 1600-tals-bibliografier visar att praktiskt taget allt svenskt 1500-talstryck excerperats och det mesta från 1600-talet (i varje fall det som är av vikt), vilket stämmer bra överens med min uppfattning av perioden från egen forskning. Det gäller biblarna, reformatorens skrifter, katekeser, psalmböcker, bönböcker, evangelier och epistlar, läkeböcker, pestböcker, mandat, stadgor, bondepraktikan och ordböcker. Därtill kommer skrifter som ej trycktes i samtiden utan senare (tänkeböcker, skotteböcker, G1R, visböcker osv.). Likaså har samlade verk av periodens förnämsta författare excerperats: Messenius, Bureus, Stiernhielm, Lucidor, Columbus, Lindschöld, Werving, Israel Holmström, Dahlstierna, fru Brenner, Runius, Frese, Triewald, Urban Hiärne, Spiegel, Swedberg, Rudbeck,

Verelius översättningar av isländska sagor, Johann Arndts skrifter. Osv. Av de 41 skrifter som Collijn förtecknar av Ericus Schroderus, som var kunglig korrektör och translator och bedrev en imponerande översättarverksamhet åren 1610-1642, har SAOB excerperat 25. Osv. Vad som ej excerperats — det gäller 1600-talet — är bl.a. ABC-böcker, enstaka visor och skrifter av mindre kända eller anonyma författare. Vad gäller ABC-böckerna är i allmänhet innehållet hämtat ur katekesen eller psalmboken, två källor som är väl excerperade och följer Bibelns språk, varför dessa källor knappast skulle tillföra ordbokens samlingar något nytt i fråga om ordförråd (eller böjnings- eller stavningsformer).

### 1700-talet

1700-talets samlade svenska bokproduktion är enorm. Den är i sin helhet ungefär 10 gånger så stor som Collijns 1600-talsbibliografi, som upptar ca 4600 titlar. Den nyligen färdiga bibliografiska databasen *Svensk Bibliografi 1700-1829*, KB, Stockholm, som grundar sig på KB:s samling av katalogiserat tryck och som är tillgänglig för alla som är anslutna till LIBRIS-systemet, uppgår till ca 40.000 tryck. I denna siffra är vissa tryckalster icke medtagna såsom årstrycket (dvs. författningstrycket, från lagar och förordningar ned till ordningsstadgor och taxor), periodica, dissertationer och program, tillfällestryck (dvs. bröllopsdikter, sorgeskväden, gravtal, namnsdagshyllningar, välgångsönskningar o.d.), sagor, skillingtryck och det kommunala förvaltningstrycket. Inbegriper man dessa tryck blir siffran 4 gånger så stor. För att få en uppfattning om hur SAOB täcker 1700-talet har jag använt *Swedish Imprints 1731-1833 (SWIM)*, ett engelskspråkigt, häftesvis utkommande arbete av nationalbibliografisk karaktär, som ges ut i Uppsala av Rolf du Rietz, Sten Lindroths *Svensk Lärdomshistoria*, Lars Svenssons *Förteckning över bibliografier rörande 1700-talets bokproduktion* och — delvis — också databasen *Svensk Bibliografi 1700-1829*. Självklart har SAOB bara excerperat en del av 1700-talets litteratur men mitt huvudintryck är att excerperingen täcker väl de viktiga och stora genrererna: vitterhet, teologi, språkvetenskap, historia, juridik, resor, geografi och topografi, juridik, teknologi, ekonomi, naturvetenskap och medicin. T.ex. har SAOB excerperat första och femte upplagan av vår första svenska nationalgeografi på modersmålet, *Inledning til geographien öfver Sverige*, författad av geografen och kanslisten Eric Tuneld 1741. De texter som Lindroth behandlar av Anders Berch, Torben Bergman, Carl Carleson, Celsius, Gadd, Klingensstierna, Linné, Nordencrantz, Rosén von Rosenstein m.fl. har excerperats av SAOB. Detsamma gäller för den topografiska litteraturen företrädd bl.a. av Broman, Broocman, Faggot, Fernow, Högström, Lars Salvius, Eric Tuneld och Oedman.

Jag har ofta förvånat mig över hur väl 1700-talet är excerperat. De äldsta SAOB-redaktionerna har haft stor litteraturkännedom men naturligtvis också använt sig av bibliografisk litteratur vid litteraturvalet. Särskilt viktiga hjälpmedel torde ha varit Lénströms *Sveriges litteratur- och konsthistoria i utkast* (1841), som innehåller över 100 sidor utförliga bibliografiska förteckningar ordnade under huvudrubriker: teologi, estetik, filosofi, botanik, zoologi, kemi, bergsvetenskap etc. Andra viktiga bibliografiska hjälpmedel torde ha varit Elmgrens *Öfversigt af Finlands litteratur* (1861-65), Lüdekes *Allgemeines schwedisches Gelehrsamkeits-Archiv unter Gustafs des dritten Regierung* (1781-1796), och Warmholtz *Bibliotheca historica Sueo-gothica* (1782-1817).

*1800- och 1900-talen*

1800- och 1900-talen är också väl representerade i SAOB:s språkprovssamling. Alla klassiska författare har excerperats från Tegnér, Geijer, Atterbom, Stagnelius, Almquist, Strindberg och Selma Lagerlöf till Hjalmar Gullberg, Edfelt, Ekelund, Ekelöf, Evert Taube m.fl. Några författare från vår egen tid som excerperats under den senaste 15-årsperioden är Stig Claesson, Sven Delblanc, Kerstin Ekman, Per Gunnar Evander, Maria Gripe, Lars Huldén, Gösta Knutsson, Ulf Lundell, Göran Palm och Bo Setterlind. Andra väl representerade genrer eller ämnesområden o.d. är biografi, botanik, ekonomi, etnografi, författning, förordning, geografi, historia, juridik, litteraturvetenskap, ordbok, periodica, samhällsvetenskap, sport, språkvetenskap, teknologi, teologi, tidningar, tidskrifter och undervisning.

**Källförteckning. Supplement II**

SAOB:s alfabetiskt uppställda tryckta källförteckning från 1939 upprättad av Sigrid von Greyerz och dess *Supplement* från 1975 upprättat av Gulli Starfelt-Brown ger inte upplysningar om vilka ämnesområden som excerperats eller vad som excerperats från en viss period utan är avsedda att förklara titelförkortningarna i de tryckta ordboksbanden. SAOB:s *Källförteckning. Supplement II*, som nyligen färdigställts, är en bibliografi över ca 1.800 texter excerperade perioden 1975-1990. Bibliografien, som lästs in i referenshanterings- eller databasprogrammet Reference Manager föreligger i datorversion (Macintosh) och är i första hand avsedd för SAOB-redaktionen men kan givetvis också vara av stort intresse för andra institutioner och forskare. Den kommer att inom kort tryckas i följande 4 delar:

- 1 en strikt alfabetisk del,
- 2 en kronologisk del,
- 3 en ämnesinriktad del och
- 4 ett ämnesordsregister.

I sökregistret kan man söka på ca 550 olika ämnen. Som sökord använder jag alla de ämnen och fackområden som finns i SAOB. Jag har dessutom lagt till ett antal sökord. Man kan också söka på författare, titel, kommentar, tryckort, tryckår, personnamn och ortnamn. Varje post eller referens har försetts med 5-7 sökord jämte underavdelningar. Den vittra litteraturen har jag försett med underavdelningarna poesi, prosa, drama, tal, svenska, översättning (från vilket språk). Söker man på "vitterhet:svenska:poesi" får man således fram all svensk originallitteratur som klassats som poesi. Söker man på "Shakespeare" så får man fram alla excerperade Shakespeare-översättningar. Söker man på "språkvetenskap:ordbok:-svensk-engelska" så får man fram svensk-engelska ordböcker. Är man intresserad av att veta var t.ex. 1700-talets litteratur trycktes utanför Stockholm, som givetvis var den dominerande tryckorten (liksom i våra dagar), så går det lätt att få fram denna uppgift. Osv. Jag har även noterat om författaren är manlig eller kvinnlig. Man kan alltså få svar på frågan hur många kvinnliga författare det fanns t.ex. under 1600-talet och vad de skrev. Sökmöjligheterna är legio. — Materialet kan sorteras med avseende på författare, titel, tryckort, utgivningsår och löpnummer.

I kommentarfältet har jag skrivit in uppgifter om eventuella original, förlagor, upplagor, översättningar, tillförlitlighet m.m., ibland också biografiska upplysningar om författaren (yrke, födelseår, födelseort o.d.). — I samband härmed vill jag nämna att docent

Sven Lange, Stockholm, använt SAOB:s källförteckning för upprättande av en databas för undersökning av svenskt skriftspråk under perioden 1526-1799. Man kan söka på författarnamn och ämnesområde men också på födelseår, födelselandskap och generation, vilket ger möjlighet att undersöka om det finns någon regional fördelning när det gäller vissa ämnesområden. De biografiska uppgifterna är mycket värdefulla för SAOB och kommer — med benäget tillmötesgående av Sven Lange — att arbetas in i SAOB:s bibliografi. De tryckta källförteckningarna har nyligen optiskt lästs in i dataprogrammet och kommer inom kort att bearbetas. Bibliografien kommer att bli en ovärderlig hjälp för ordboksredaktionen när det gäller att snabbt upptäcka luckor i excerperingen. Det blir också lätt att uppdatera databasen.

Fig. 3 Sökning på några slumpvis valda sökord i databasen Källförteckning. Supplement II.

sökord	antal källor	tryckår
bin	1	1969
dator	4	1973, 1979, 1980, 1982
fiske	7	1754, 1914, 1935, 1941, 1947, 1952, 1954
schack	6	1839, 1869, 1902, 1908, 1972, 1979

De undersökta ämnesområdena är förutom "bin" väl representerade, även kronologiskt.

### Optisk inläsning av SAOB

Artiklarna i SAOB innehåller en mångfald uppgifter av skilda slag. Läsaren får information om uttal, ordklass, böjning, sido- eller variantformer, stavning, etymologi, bruklighet, betydelse, äldsta (och för ord med inskränkt bruklighet också yngsta) belägg, sammansättningar, avledningar, och i fråga om verb särskilda förbindelser osv. Vid Språkdata i Göteborg pågår sedan ca 10 år tillbaka ett projekt med syfte att upprätta en databas över de tryckta banden av SAOB. I dag är artiklarna i banden 1-28 inlästa och kontrollarbete pågår. Databasen, som kommer att ge en mängd sökmöjligheter — kronologiska, morfologiska, stilistiska, etymologiska osv. — kommer att bli ett värdefullt hjälpmedel för ordboksredaktionen och annan lingvistisk forskning. Ordboksredaktionen kommer inom kort att kunna utnyttja databasen — i varje fall delar av den. Det kommer också att vara fullt möjligt att koppla databasen till den *Bibliografi över SAOB:s källmaterial* som utarbetas vid redaktionen i Lund. Genom att med Bibliografins hjälp spåra texter från ett visst ämnesområde — t.ex. juridik — kan man med databasens hjälp komma åt källhänvisningarna och därmed uppslagsorden och på så sätt upprätta en förteckning över ordförrådet i juridiskt fackspråk från t.ex. 1700-talet. Tanken svindlar inför de möjligheter som yppar sig.

### SAOB:s språkprovssamling tillgänglig för forskning

Ett stort antal forskare från Norden och andra länder har under årens lopp besökt ordboksredaktionen i Lund för att undersöka och excerpera SAOB:s väldiga material, varav endast ca 10-15 % redovisas i tryck. Excerptssamlingen utnyttjas inte bara av språkforskare utan

också av litteraturhistoriker, teologer, historiker, folklivsforskare, jurister osv. Särskilt för ordforskare erbjuder språkprovssamlingen ett oskattbart och oundgängligt material. Erik Terner, t.ex., utvidgade sin ordboksartikel om räkneordet *En* till en doktorsavhandling (1922), och Emil Olson gav ut sin ordboksartikel *Den* som en egen skrift (1913). Elof Hellquist fick — har det berättats — idén till sin monumentala *Svensk etymologisk ordbok* (1920-22) genom sitt arbete som redaktör vid SAOB, där han var anställd 1893-1904. Och — för att avsluta i vår egen tid — för något år sedan har Elsie Wijk-Anderssons doktorsavhandling utkommit, *Bara i fokus*, som är en semantisk-syntaktisk studie av adverbet *bara* och dess ekvivalenter (*allena, allenast, blott, enbart, endast, uteslutande*) i nysvenskt skriftspråk, ett ämne som knappast skulle kunna ha behandlats utan SAOB och dess materialsamling.

## Noter

1. Boktitlar o.d. förkortas enligt SAOB:s principer. Se Litteratur och förkortningar sid. ??

## Litteratur och förkortningar

- Allén, Sture. 1980. Hur djupt går SAOB? I: *Nysvenska Studier* 60: 119-129. Upps.
- Allén, Sture. 1986. Inför tredje seklet. I: Allén, Sture, Loman, Bengt, Sigurd, Bengt, *Svenska Akademien och svenska språket. Tre studier.* 233-273. Sthm.
- Collijn, Isak. 1927-38. *Sveriges bibliografi intill år 1600.* 1-3. Upps.
- Collijn, Isak. 1942-44. *Sveriges bibliografi. 1600-talet.* 1-2. Upps.
- Ekbo, Sven. 1965. Historik, organisation och arbetsmetod. I: Ekbo, Sven, och Loman, Bengt, *Vägledning till Svenska Akademiens ordbok.* NSS 32. Sthm.
- Ekbo, Sven. 1965. Ordboksartikelns uppbyggnad. I: Ekbo, Sven, och Loman, Bengt, *Vägledning till Svenska Akademiens ordbok.* NSS 32. Sthm.
- Elmgren, Sven Gabriel. 1861-65. *Öfversigt af Finlands litteratur* [1-2]. Hfors.
- Falck-Kjällqvist, Birgit. 1987. Om Svenska Akademiens ordbok. I: *Språkbruk* 1. 16-20. Hfors.
- GIR = Konung Gustaf den förstes registratur.* 1-29. 1861-1916. Sthm.
- Hast, Sture. 1983. Om SAOB:s material. I: *Nysvenska Studier* 63. 159-192. Upps.
- Jonsson, Hans. 1981. Ekeblad och SAOB. I: *Nysvenska Studier* 61. 134-138. Upps.
- Jonsson, Hans. 1983. Om historisk lexikografi. I: *Nysvenska Studier* 63. 128-158. Upps.
- Jonsson, Hans. 1990. Svenska Akademiens Ordbok (SAOB). I: *Wörterbücher .. Ein internationales Handbuch zur Lexikographie...* Edited by .. Franz Josef Hausmann et.al. Second Volume. 1940-1943. Berlin • New York.
- KrigVAH = Kongl. krigs vettenskaps akademiens handlingar.* 1805 ff. Sthm.
- Krig VAT = Kongl. krigs-vetenskaps akademiens tidskrift.* 1833 ff. Sthm.
- Källförteckning* [till SAOB]. [1898]. Lund.
- Källförteckning* [till SAOB]. [1939]. Lund.
- Källförteckning. Supplement* [till SAOB]. [1975]. Lund.
- Källförteckning. Supplement II* [till SAOB]. 1990. Lund.
- Lange, Sven. 1986. Databas för undersökning av svenskt skriftspråk perioden 1526 till 1799  
I: *Datorstödd språkforskning. Fyra uppsatser om lånord, växtnamn, språk- och stilhistoria samt ordgeografi.* MINS 22. 25-32. Sthm.
- Lénström, Carl Julius. 1841. *Sveriges litteratur- och konsthistoria i utkast.* Ups.

- Lindroth, Sten. 1975. *Svensk lärdomshistoria. Frihetstiden*. Sthm.
- Lindroth, Sten. 1975. *Svensk lärdomshistoria. Gustavianska tiden*. Sthm.
- Lüdeke, Christoph Wilhelm. 1781-1796. *Allgemeines schwedisches Gelehrsamkeits-Archiv unter Gustafs des Dritten Regierung*. 1-17. Leipzig.
- Olson, Emil. 1913. *Studier över pronominet den i nysvenskan*. I: Lunds universitets årsskrift. NF. Avd. 1. Band 9. Nr 3. Lund.
- OxBr.* = Rikskansleren Axel Oxenstiernas skrifter och brevexling. Afd. 2: 1 ff. 1888 ff. Sthm.
- PH* = Utdrag utur alle ifrån .. 1718 utkomne publique handlingar. 1-15. 1742-1829. Sthm.
- Rietz, Rolf E. du. 1977 ff. *Swedish Imprints 1731-1833 [SWIM]. A Retrospective National Bibliography ..* Upps.
- Rydstedt, Rudolf. 1988. Creating a Lexical Database from a Dictionary. I: *Studies in Computer-Aided Lexicology*. 228-267. Gbg.
- SAOB* = Ordbok öfver svenska språket. 1898 ff. Lund.
- SAOL* = Svenska Akademiens ordlista över svenska språket. 11 uppl. 1986. Sthm.
- Starfelt, Gulli. 1975. Något om citatkontrollen vid Svenska akademiens ordboksredaktion. I: *Festskrift till Sven Ekbo*. 122-128. Sthm.
- Svensk Bibliografi 1700-1829*. [Databas upprättad av Sune Lindqvist, Gunilla Jonsson och Ylva Tjeder. Kungliga Biblioteket, Sthm.]
- Svensson, Lars. 1988. En bibliografi över SAOB.s källmaterial på persondator. I: *Sagt och skrivet. Festskrift till David Kornhall*. 155-157. Lund.
- Svensson, Lars. 1985. *Förteckning över bibliografier rörande 1700-talets bokproduktion. Ett urval*. Nordlund. 8 [stencilupplaga]. Lund.
- Svensson, Lars. 1985. Om 1700-talets bokproduktion. I: Ulf Teleman, red. *Det offentliga språkbruket och dess villkor i Sverige under 1700-talet*. Nordlund 7 [stencilupplaga]. 61-78. Lund.
- Terner, Erik. 1922. *Studier över räkneordet en och dess .. användning .. i nysvenskan*. Upps.
- Warmholtz, Carl Gustaf. 1782-1817. *Bibliotheca historica Sueo-gothica; eller förteckning uppå så väl trykte, som handskrifne böcker, ..* 1-15. Sthm och Ups.
- Vifot, Eva. 1975. Hur ordboksartiklar förbereds. I: *Kring en ordbok. Festskrift till Sven Ekbo*. 195-202. Sthm.
- Wijk-Andersson, Elsie. 1991. *Bara i fokus. En semantisk-syntaktisk studie av bara och dess ekvivalenter i nysvenskt skriftspråk*. Upps.
- ÅrsbVetA* = Årsberättelse om vetenskapernas framsteg, afgifne af kongl. vetenskaps--academiens embetsmän. 1822-59. Sthm.

## En ordbog og dens informanter

På Orðabók Háskólans i Reykjavík, Islands leksikografiske institut, har man i næsten femti år indsamlet materiale til en historisk ordbog<sup>1</sup>. Jeg vil fortælle om ét led i dette indsamlingsarbejde, det vil sige hvordan Ordbogen har draget nytte af den almene sproglige interesse i landet ved at rette spørgsmål til folk angående ord der ikke er fundet i skriftlige kilder eller der mangler nærmere oplysninger om. Jeg vil dog først give et kort overblik over den virksomhed der er foregået på Ordbogen i de forløbne årtier.

Planer om en historisk islandsk ordbog blev først fremsat i begyndelsen af dette århundrede, og da ville man stile mod en ordbog som dækkede islandsk sprog fra de ældste kilder og op til sidste århundredskifte. Man startede arbejdet, men på grund af svigtende bevillinger måtte man snart give op. Idéen blev dog holdt i live, og i begyndelsen af firtiårene gik man på ny i gang med arbejdet med en historisk ordbog, som ifølge nye planer skulle omfatte perioden fra 1540 til nutiden. Sproghistorisk set er den nye tidsgrænse bagud tilfældig, men 1540 er året for udgivelsen af den første bevarede trykte bog på islandsk, Det ny testamente, og en forklaring på at man nu satte denne grænse var den at man da i København var begyndt at forberede et andet historisk ordbogsværk, som skulle omfatte norrøn prosa, og heri indgik islandsk op til 1540.

I de forløbne år har indsamlingsarbejdet bestået i excerpering af skrifter fra denne periode, især trykte bøger, og man kan sige at man allerede har excerperet de vigtigste trykte kilder op til slutningen af forrige århundrede foruden et betydeligt materiale fra det tyvende århundrede.

Imidlertid forholder det sig sådan, at de trykte islandske bøger fra det 16., 17. og op i det 18. århundrede først og fremmest indeholder kirkelig og anden opbyggelig litteratur, og derfor har det været nødvendigt at foretage en betydelig excerpering på grundlag af et bredere tekstudvalg, det vil sige fra håndskrifter, for at udfylde lakuner. Blandt andet har man søgt at inkorporere alle håndskrevne ordbøger og ordsamlinger. Samlingen fra trykte kilder og håndskrifter består nu af over 2 millioner ordbogssedler, og i de senere år har man arbejdet med at indtaste dette materiale, således at det bliver maskinelt tilgængeligt.

Men en historisk ordbog må give et så fyldigt og detaljeret billede som muligt af sproget i den periode den omfatter, og skriftsproget alene dækker aldrig det levende sprog til fulde. Til trods for omfattende bogproduktion inden for mange forskellige emner i løbet af de sidste to århundreder, dækker disse kilder aldrig hele det ordforråd som er eller har været forbundet med forskellige arbejdsmetoder og kulturelle foreteelser. Mange slags udtryksmåder bliver således sjældent eller aldrig skriftfæstet, i hvert fald i trykte bøger, og et stort antal ord og ordforbindelser, særlige betydninger og talemåder, som hører hjemme i en udførlig

---

<sup>1</sup> Senere omtalt som Ordbogen.



historisk ordbog, er således i fare for at gå tabt. Derfor er det nødvendigt for leksikografer også at søge materiale i det talte sprog, hvis de skal være i stand til at foretage en historisk beskrivelse af ordforrådet, sprogets udvikling og betydningsændringer.

Det er en lang islandsk tradition for at interessere sig for talesproget og optegne dialektale ejendommeligheder. I begyndelsen af det 18. århundrede indsamlede håndskriftsamleren og forskeren Árni Magnússon fra forskellige landsdele ord og talemåder, som var ham selv fremmede og som han mente var karakteristiske for de pågældende landsdele. Hans amanuensis, Jón Ólafsson fra Grunnavík, som gennem et langt liv arbejdede med et stort anlagt ordbogsværk, optegnede meget fra sin samtids talesprog og bestræbte sig på at oplyse om kilderne til de optegnede ord og ordforbindelser.

Den danske sprogforsker Rasmus Kristian Rask foretog en rejse i det østlige Island tidligt i det 19. århundrede og optegnede på sin rejse ord som han anså for kendetegnende for den østlandske dialekt.

Den islandske filolog Björn M. Ólsen, som senere blev professor og den første rektor ved Islands universitet, viste stor interesse for talesproget og foretog optegningsrejser i landet i årene 1884-93, det meste af tiden med støtte fra Carlsbergfondet.

Man kan sige at det første forsøg på en organiseret indsamling fra talesproget begyndte i 1917 da forfatteren Þórbergur Þórðarson fik en bevilling fra Altinget til indsamlingsarbejde. Han fik årlige bevillinger indtil 1924, da de blev sløjfet og dette arbejde ophørte. Þórbergur lagde stor vægt på at få kontakt med meddelere så mange steder som muligt over hele landet og opbygge et samarbejde mellem frivillige. Hans formål var en systematisk indsamling i hver syssel af ordforråd der var knyttet til bestemte arbejdsområder og -metoder fremfor tilfældige optegnelser af enkelte ord. I en periode forsøgte han også at udvide indsamlingsarbejdet gennem en fast avisspalte, og det lader til at en del folk har reageret herpå. Da bevillingerne til ham ophørte, opgav han at have omkring 300 informanter.

Alle de samlinger som jeg her har omtalt er nu helt eller delvis i Ordbogens varetægt. Dog ser det ud til at noget af Þórbergurs materiale er gået tabt. På Ordbogen opbevares endvidere en del mindre samlinger som jeg ikke har nævnt. Dette materiale er kun delvis blevet behandlet endnu, men det er tilgængeligt for alle som vil undersøge ordforrådets historie eller dialektale udtryk i enkelte landsdele. Samtidigt er disse kilder et uvurderligt bidrag til det indsamlingsarbejde som foregår på Ordbogen. Uden dem ville en del udtryk som kendes fra skriftsproget være svære at forklare, og en del oplysninger som disse samlinger har bevaret ville ellers være gået tabt.

Ordbogens stab har uden tvivl fra starten gjort sig klart hvor vigtigt det var at indsamle materiale fra talesproget, men det første årti har man dog ikke taget noget initiativ på dette felt, nok især som følge af hvor fåtallige man var og hvor meget andet arbejde der skulle varetages.

Midt i halvtredserne blev Ordbogens redaktører opfordret til at tage sig af regelmæssige radioudsendelser om islandsk sprog, indslag som radioen flere år forud havde haft med forskellige løse medarbejdere. Der var her tale om oplysende programmer med vejledning med hensyn til sprogbrug i skrift og tale. Under ordbogsfolkernes redaktion udviklede disse programmer sig efterhånden til at blive en blandning af oplysning og indsamling. Redaktørerne havde da siddet med excerperingsarbejde i flere år og kunne formidle mange interessante oplysninger til lytterne, blandt andet om ord og udtryk som tilsyneladende var begrænset til bestemte landsdele eller mindre områder. Samtidig kunne de drage nytte af lytternes reaktioner ved deres indsamlingsarbejde, idet lytterne begyndte at formidle oplysninger til Ordbogen. Disse udsendelser blev meget aflyttet, og Ordbogen opnåede hurtigt kontakt med mange gode informanter, og nogle af disse har indsendt

oplysninger i årtier. Enkelte af de første informanter havde i sin tid bidraget til Þórbergur Þórðarsons indsamlingsarbejde.

Ordbogsfolkene tog ikke bare imod svar på deres forespørgsler. De meddelte lytterne oplysninger som de modtog, i det omfang svarene gav anledning dertil, og det har man holdt fast ved siden. Og lytterne svarer ikke bare spørgsmål. De stiller ofte selv spørgsmål om ord og udtryk, om oprindelse, brug og betydning, og i det omfang tiden tillader det benyttes radioudsendelserne også til at give svar på sådanne spørgsmål på grundlag af de oplysninger man allerede sidder inde med. I andre tilfælde må spørgsmålet sendes videre til andre lyttere i håb om reaktioner fra dem. Selv om Ordbogens hovedformål med udsendelserne naturligvis er at indsamle materiale, er det på denne måde lykkedes at opholde en tosidet kommunikation med folk fra forskellige samfundslag rundt omkring i landet. Det er nu tale om et nyt indslag hver uge i vintermånederne på program 1, og hver udsendelse gentages et par dage senere. Længden har været noget varierende, som regel tyve minutter, men i nogle perioder ti eller femten minutter. I den første tid modtog Ordbogen næsten altid lytternes svar i skriftlig form, men efterhånden som telefonsystemet blev udbygget og det blev nemmere og billigere at telefonere fra landet, blev det mere og mere almindeligt at lytterne i fortsættelse af en udsendelse ringede den pågældende ordbogsredaktør op - enten på Ordbogen eller privat. Selv om telefonsamtalerne unægtelig tager tid - bl.a fordi nogle informanter også trænger til en personlig sludder - så er telefonsamtalerne i visse henseender mere givende end brevene, idet de giver lejlighed til at spørge nærmere om de behandlede ords betydning og samtidig eventuelt om andre ord inden for det samme område. Man kan også benytte lejligheden til at spørge om ord som man har modtaget oplysninger om fra andre personer i nabolaget.

Foruden svar på direkte forespørgsler har flere informanter sendt Ordbogen kortere eller længere lister over sjældne ord som de har hørt og noteret hos sig, måske gennem en længere periode, også lister over ord som de har stødt på under læsning og som har været dem fremmede. Dagbøger og andre skrevne kilder har vi også modtaget til excerpering. Alt dette materiale bliver gennemgået på Ordbogen, og vi får på denne måde tit oplysninger om ord som i forvejen ikke eller kun i ringe grad er belagt i vores seddelsamlinger. Disse oplysninger kan i visse tilfælde give anledning til omtale i radioudsendelserne, og ofte har det vist sig at ord eller udtryk, som man tidligere havde troet var meget isolerede, faktisk hørte til det levende sprog i et større område.

De oplysninger som Ordbogen modtager fra lytterne bliver omhyggeligt registreret på sedler. Foruden selve opslagsordet anføres informantens definition af ordets betydning så præcis som muligt. På sedlen noteres informantens navn og bopæl og hans eller hendes kilde. Til slut oplyses om informationen er givet per brev eller i en samtale. Alle indsendte breve arkiveres, og de udgør efterhånden en stor samling.

Sedler som indeholder oplysninger fra det talte sprog holdes for sig på Ordbogen. Det vil sige at de ikke ordnes ind imellem sedler der bygger på excerpering af skriftlige kilder, idet disse to samlinger er væsensforskellige. Eksemplerne fra de skriftlige kilder viser ordenes brug på en bestemt tid i en bestemt kontekst, medens eksemplerne fra det talte sprog snarere ledsages af informantens oplysninger om betydning og udbredelse. I Ordbogens talesprogs-samling er der nu over 190 tusinde sedler, men man har ikke optalt hvor mange opslagsord de dækker.

I løbet af af de tre-fire årtier som Ordbogens redaktører har stået for disse sprogindslag har Ordbogen, som før nævnt, fået kontakt med et stort antal informanter i alle aldersgrupper. For fem år siden lavede man en statistik over Ordbogens informanter før og nu, som indeholdt alle de oplysninger om vedkommende, som vi var i besiddelse af. Antallet viste sig til at være næsten 2400, heraf knap 1500 mandlige, og den geografiske spredning var meget god. Nogle

af disse informanter har regelmæssigt eller med mellemrum formidlet Ordbogen oplysninger gennem det meste af denne periode, andre har optrådt mere sporadisk og en del er naturligvis i tidens løb afgået ved døden. Hvert år dukker der imidlertid nye informanter op, således at det samlede antal er vokset.

Måske har det spørgsmål meldt sig om oplysninger indhentet på denne måde er pålidelige. Det kan man i de fleste tilfælde svare ja til. Den nære kontakt, som der normalt er mellem informanterne og Ordbogen, gør at de ordbogsredaktører som tager sig af udsendelserne og modtager svarene lærer de enkelte informanter personligt at kende og er derved i stand til at evaluere de oplysninger de afgiver. Ved en gennemgang af talesprogs-samlingen for nogle år siden viste det sig, som det var at forvente, at den indeholdt en del hapax legomena som krævede nærmere undersøgelse. Man spurgte i radioen om hovedparten af disse ord, men kun i få tilfælde modtog vi en bekræftelse af deres eksistens. Man kan således regne med at nogle af disse ord er opstået ved at informanten har hørt forkert eller misforstået. Men dette er et problem som ikke er specielt for disse oplysninger, idet tilsvarende problemer opstår når skriftlige kilder excerperes, og sådanne isolerede former, som ikke får nogen støtte, bliver lagt til side.

Som det er fremgået af det der allerede er sagt, består udsendelserne for det meste af spørgsmål og svar angående enkelte ord og ordforbindelser. Spørgsmålene giver lytterne imidlertid anledning til også at nævne andre ord inden for det samme område, eksempelvis husdyrfarver, høhøst og forskellige slags redskaber, således at betydelige oplysninger foreligger nu angående enkelte fænomener. Det er lykkedes at indsamle et betydeligt ordforråd der er forbundet med dele af dagliglivet som der ikke er skrevet meget om. Dette gælder for eksempel vejforhold, forskellig slags håndarbejde som dels blev udført af mænd og dels af kvinder, tøj, madlavning og det daglige arbejde i det hele taget både til sø og til lands. Det drejer sig her naturligvis ikke mindst om ordforråd der er forbundet med aflagte arbejds-metoder, og radioudsendelserne har bidraget væsentlig til at en betydelig del af dette ordforråd ikke er gået i glemmebogen.

Det er dog ikke gennem radioen foretaget en systematisk indsamling af ordforråd der er forbundet med bestemte arbejdsområder eller -metoder, og det materiale der er blevet indsamlet er ikke blevet klassificeret emnemæssigt. De enkelte redaktører har ganske vist ofte i en periode valgt først og fremmest at behandle specielle områder og har høstet meget materiale, men lytternes reaktioner har været meget forskellige af omfang. Ændrede arbejdsmetoder medfører at ordforråd, der har været forbundet med det gamle, hurtigt bliver fjernet og kun huskes af få, og i de tilfælde, hvor det drejer sig om arbejde der kun er blevet udført af et fåtal mennesker, er lytternes reaktioner ofte minimale. Derfor er det vigtigt at sammenflette det almene og det mere specifikke i udsendelserne. Så snart man kommer ind på de mere almindelige områder, får vi som regel mange svar og der dukker gerne nye hjemmelmænd og -kvinder op, som tidligere havde ment at de ikke havde så meget at formidle, men som kan vise sig at være meget vigtige informanter inden for forskelligartede emner.

Jeg har nu givet en kort oversigt over tidligere indsamling fra det talte sprog på Island og beskrevet den metode som Islands leksikografiske institut har benyttet sig af for at forøge sine samlinger. Udsendelserne har haft en trofast lytterskare, og aflytningen og interessen forekommer ikke at være blevet væsentlig formindsket til trods for at Statsradiofonis program 1, som i starten var landets eneste ærtermedie, efterhånden har fået konkurrence af flere radioprogrammer og to fjernsynskanaler. Men hvilket udbytte har Ordbogen selv haft af sine udsendelser i årenes løb?

Der har været en betydelig forbindelse mellem excerpering fra skriftlige kilder og indsamling fra det talte sprog. Ved excerpering fra bøger dukker der ord og ordforbindelser op, hvis nøjagtige betydning ikke fremgår af konteksten og som man ikke i forvejen har belæg på. Man kan slå forgæves op i eksisterende ordbøger, og da er det oplagt at søge hjælp hos Ordbogens informanter. I mange tilfælde er det at den vej lykkedes at indhente gode beskrivelser af vedkommende udtryks betydning og anvendelse.

Talesprogssamlingen er i høj grad blevet brugt ved indsamling af materiale til forskellige undersøgelser inden for etymologi, leksikografi, dialektgeografi og morfologi. Det har vist sig at samlingen rummer betydelige oplysninger om udtryk som ikke kendes fra trykte bøger. Den indeholder mange eksempler på talemåder og ordsprog som ikke findes i de bøger der er blevet skrevet om disse emner. Der kan man også finde forskellige bøjningsformer og udtaleformer som i visse tilfælde er lokalt begrænsede og kan bidrage til at kaste lys over den sproglige udvikling i bestemte dialekter.

I ældre ordbøger findes der tit oplysninger om at bestemte ord og ordforbindelser er lokale, men ofte har forespørgsler til lytterne afsløret at de er kendt over et langt større område, i visse tilfælde over hele landet. Forklaringen er at tidligere ordbogsforfattere har holdt sig til de begrænsede oplysninger de havde, og de kunne stamme fra én bestemt egn. De havde ingen mulighed for at undersøge den faktiske udbredelse på den måde som vi nu kan.

Der findes eksempler på at udtryk i talesprogssamlingen har kunnet bidrage til en sikrere forståelse af tvivlsomme steder i det gamle sprog, og det viser sig også nogle gange at ord, som man kun havde belæg på i middelaldersproget, stadig hører til det levende ordforråd visse steder i landet.

Af det foregående skulle det være klart at indsamling af materiale fra det talte sprog må gå hånd i hånd med anden indsamling til en historisk ordbog. En historisk ordbogs seddelsamlinger tjener ikke bare det formål at være ordbogsmateriale, idet de også undersøger en uundværlig kilde ved de fleste undersøgelser af sproget og dets udvikling. Det institut, som har til formål at udarbejde en historisk ordbog, er således samtidig et alsidigt leksikografisk institut.

Det er nødvendigt at gøre talesprogssamlingen maskinelt tilgængeligt i en database som ville gøre den lettere at anvende til forskelligartede formål. Man burde således for eksempel anføre nøjagtigt den geografiske placering, således at man kunne kalde frem ord fra bestemte landsdele. Man burde endvidere markere talemåder, ordsprog, bøjningsformer og udtaleejendommeligheder. Det ville også være ønskeligt at klassificere materialet, således at man kunne kalde frem ordforråd inden for bestemte områder såsom folketro, folkekultur, arbejdsmetoder og værktøj.

Det står klart at radioprogrammerne alene ikke er tilstrækkelige til at opsamle alt det ordforråd som efterhånden går tabt. Det ville være nødvendigt at indsamle systematisk gennem udarbejdelse og udsendelse af ordlister og ved besøg hos informanterne, men en sådan indsamling kræver grundig forberedelse, og den vil være tidskrævende og bekostelig. I den forbindelse skal det dog nævnes, at Nationalmuseets afdeling for folkekultur i de sidste årtier har gennem spørgelister indsamlet stort materiale på forskellige områder, og det materiale gøres nu efterhånden maskinelt tilgængeligt og dermed lettere anvendeligt for Ordbogen til excerpering; men denne indsamling er naturligvis ikke primært foretaget med et leksikografisk sigte. Vores radioprogrammer har til trods for deres begrænsninger gjort stor nytte, men denne metode er muligvis bedst egnet i et lille og forholdsvis homogent sprogsamfund. Metoden er også blevet anvendt på Færøerne, hvor den også har vundet popularitet. Og den kunne måske også tilskynde andre ordbogsredaktører til et øget samarbejde med almindelige sprogbrugere.

Helle Degnbol, Guðrún Ása Grímsdóttir, Bent Chr. Jacobsen, Jette Knudsen, Eva Rode & Christopher Sanders

## En ordbog er en database

Under produktionen af *Ordbog over det norrøne prosasprog: Registre* (første bind af Den arnamagnæanske kommissions ordbog i 12 bind, udkommet i slutningen af 1989) lagdes alle oplysninger om ordbogens kildemateriale (håndskrifter, udgaver, norrøne og udenlandske) i database. Fra databaserne udvalgte redaktionen, hvilke data pr. automatik skulle bringes i registerbindet og under hvilken præsentationsform (typografi, layout). Det traditionelle stadium i en bogs tilblivelsesproces med udfærdigelse af "tekstbehandlet" manuskript og efterfølgende korrekturlæsning kunne således overspringes, og arbejdet kunne koncentreres om bogens indhold.

For halvandet år siden udkom det første af tolv ordbogsbind: *Ordbog over det norrøne prosasprog: Registre / A Dictionary of Old Norse Prose: Indices*. Bindet præsenterer ordbogens korpus, det norrøne (dvs. det middelalderlige islandsk-norske) prosasprog, i form af et signaturregister med oversigt over benyttede udgaver og håndskrifter (selve "nøglen"), et register over benyttede middelalderhåndskrifter med dateringer og litteraturhenvisninger, samt et register over benyttet litteratur, første del af ordbogens akkumulerende bibliografi.

For en kort beskrivelse af ordbogsprojektet henvises til *Lexikonord: Leksikografi i Norden*, Nordisk språksekretariats rapporter 14 (Oslo 1991:67-71).

Specielt om baggrunden for registerbindet og dets produktion kan man læse i The Arnamagnæan Institute & Dictionary: *Bulletin 17: 1988-1989* (København 1990:23-27).

Prøver på ordbogsartikler ses i *Ordbog over det norrøne prosasprog: Prøvehæfte / A Dictionary of Old Norse Prose: Prospectus*, Kbh. 1983.

Al virksomhed er nu koncentreret om produktion af artikler til det næste ordbogsbind. I det følgende præsenteres glimtvis hvordan vi rent teknisk arbejder med at lave "egentlig" ordbog: i princippet som ved udarbejdelsen af registerbindet, dvs. i database fra første færd og så nær hen til den endelige bogproduktion som muligt, i et system, vi selv har opbygget og til stadighed udvikler.

### Database

Hvad vi forstår ved en database, er en samling oplysninger indlagt i en fast struktur. Ordbogens database er emnemæssigt opdelt i enheder kaldet filer.

Centralt står ord-/hovedfilen med lemma, ordtypebestemmelse (til udskillelse af fx. person- og stednavne og poetisk ordforråd), grammatisk bestemmelse og oplysning om ordenes bøjning og formvariation (af denne fil dannes *artikelens hoved*), citatfilen og definitionsfilen (af hvilke dannes *artikelens krop*).

Perifert står filer med komposita-forled og med oplysninger om ordenes behandling i anden litteratur, herunder ordbøger (af hvilke filer dannes *artikelens hale*, afsnittene Comp., Litt. og Gloss.).

Ligeledes perifert står en fil med oplysning om de til grund for signaturerne liggende udgaver og håndskrifter, en værkfil, samt en håndskriftfil med oplysning om datering, proveniens mm. (Disse filer danner grundstammen i *registerbindet*.)

Endelig hører der til hele systemet en bibliografisk fil (efter hvilken den akkumulerende *bibliografi* trykkes).

Hver af disse filer er opdelt i mange, mindre felter, hvert felt rummende sin information. Således har alle oplysninger, store som små, deres plads i en ordnet struktur. Med sin konsistente arkitektur er systemet fleksibelt og giver mulighed for datavalidering på højt niveau, fx. kontrol af signaturer; udnyttelse af allerede lagrede informationer, fx. datering af citatmaterialet; beskuelse af materialet under forskellige synsvinkler; rettelser og omflytninger; vilkårligt bestemte udskrifter; automatisk registrering af hvert ords status i redaktionsprocessen. Alt bestemmes og udføres direkte af redaktørerne.

### Citatmateriale

Kernen i ordbogsartikelen og dens hele udgangspunkt er citaterne, og den mest fundamentale enhed i databasen er den fil, der indeholder ordbogens citatmateriale.

ONP's seddelsamling rummer mange forskellige citattyper: enkle citater med kun én tekst (hentet enten i en udgave eller direkte i et håndskrift); citater med en eller flere varianter, repræsenterende forskellige versioner af samme tekst; citater med udenlandsk parallel; citater med udgivers rettelse; citater med udgivers kommentar; citater med ordbogsredaktørs kommentar, evt. i form af norrøn hjælpetekst eller i form af oversættelse til dansk/engelsk. Ofte er disse citattyper kombineret, og der er derfor i databasen brug for mange felter svarende til alle de forskellige lag og typer af oplysninger. Citatfilen har således henvend 100 felter.

Databasen hjælper på forskellig vis til rationel behandling af materialet, og der skal her gives et enkelt eksempel (et mindre pilotprojekt): En type sedler, der traditionelt vil koste en del arbejde, er dem, der rummer citater afskrevet direkte fra håndskrifterne, idet de pågældende tekster endnu ikke er udgivet. Ordsedlerne indtastes, som de optræder i skufferne, dvs. med citater fra mange forskellige håndskrifter i tilfældig rækkefølge. Herpå sorteres de efter håndskrift og håndskriftside. Man har nu et praktisk overblik over materialet og kan sammenligne citat og originaltekst håndskrift for håndskrift og side for side, man kan nemt indsætte linjenummer, som er det der oftest mangler, kontrollere ortografien, tilføje mere kontekst osv. På denne måde er det kun nødvendigt at konsultere det samme håndskrift én gang.

På samme måde kan databasen sikre rationel korrekturlæsning af ordbogens mange citater imod udgaverne.

Når redaktøren modtager citaterne med henblik på redigering, er de blevet indtastet af assistenter. I samme proces er signaturerne kontrolleret og revideret og visse systematiske fejl påvist og udryddet.

For såkaldt mindre ords vedkommende indtastes alle ordsedler. Inden for bogstavet a er der ca. 40.000 citater fordelt på ca. 4.000 ord; 83 % af ordene er belagt med mindre end 9 eksempler, og hele 96,5 % af ordene er på under 50 belæg. De indtastes rub og stub. Inden for de såkaldt større ord, de 3,5 % af ordene, der er belagt med mere end 50 (enkelte med

over 1000) eksempler, sorteres der fra inden indtastning og redigering. Selv om det er ærgerligt at måtte undvære en del af citaterne i databasen, ses det klart, at frasorteringen betaler sig, da disse få større ord udgør hele 55 % af seddelmaterialet.

Citaterne kan redaktøren få i hænde (dvs. op på sin skærm eller printet ud i et hvilket som helst ønsket format, fx. seddelformat) sorteret efter eget valg: altid naturligvis efter ord (lemma), men dernæst fx. alfabetisk efter navnene på de værker, der er citeret (det kan være praktisk ved meget store ord) eller kronologisk efter de tilgrundliggende håndskrifers alder. Håndskriftdatering er automatisk indført i filen og påført udskriften, og ordets ældste citat er automatisk markeret, med et mærke, der følger citatet hele vejen ud i den trykte ordbog. Det vil altid være værdifuldt for redaktøren at få dette kronologiske overblik over ordet. Det kan også være frugtbart ved databasemæssig sortering at få samlet alle citater, til hvilke der er udenlandske paralleller, ligesom citaterne kan vælges sorteret efter værkgenre.

## Redigering

Redigeringen er den mest udfordrende del af ordbogsproduktionen, både i sig selv og i forhold til databasesystemet.

Redigeringsarbejdet i alle dets faser foregår (eller kan foregå) i database. Mest aktivt redigeres der på de centrale filer: ord-/hoved-, definitions- og citatfilerne, som tilsammen skal udgøre artikelens hoved og krop.

Redaktøren kan i princippet begynde redigeringen hvorsomhelst i systemet, men han begynder gerne i *citatfilen*, ligesom man ved traditionel redigering oftest tager fat på citatmaterialet først og sorterer sedler. Ved hjælp af særlige redigeringsfelter foretager redaktøren den første analyse af citaterne og tager for hvert citats vedkommende stilling til ordets/citatets betydning og brug, indsætter evt. kommentar til eller særoversættelse i citatet, ser på ordets syntaktiske stilling og eventuelle indgåen i en ordforbindelse af løsere eller fastere karakter, samt danner sig et overblik over norrøne varianter og udenlandske paralleller.

Det er på dette stadium naturligt at formulere definitioner og beslutte sig for artikelens inddeling og opstilling, kort sagt lægge data i *definitionsfilen*. I citatfilens særlige indekseringsfelter markeres det, hvilke citater hører til hvilken betydningsafdeling (evt. syntaktisk afdeling eller afdeling med ordforbindelse), hvilke belæg skal citeres og i hvilken rækkefølge, hvilke blot bringes som 'nøgen henvisning' (signatur uden citat), og hvilke ikke skal med overhovedet. I det integrerede databasesystem ruller definitioner og citater koordineret hen over skærmen. Indekseringsfelterne gør det let og smidigt at ændre i artikelen, idet den så at sige sorterer sig selv. Fx. vil et citat ved ændring af dets betydningsnummer automatisk flytte plads, og det er tilsvarende nemt at ordne citaternes rækkefølge inden for de forskellige afdelinger. I hver afdeling følger efter citaterne et udvalg af nøgne henvisninger i kronologisk orden.

I databasesystemet får man et klart overblik over ordets ortografiske variation og dets bøjningsformer og kan dermed færdigudfylde *ord-/hovedfilen*.

Når centralfilerne er færdigbearbejdet, er artikelens hoved og krop formet. De afsnit, der udgør artikelens hale med stof fra både de centrale og de perifære filer, genereres næsten automatisk, dvs. med meget lidt manuel indblanding fra redaktørhold. Latinske, franske, tyske, engelske og danske paralleller indhentes fra citatfilen, sammensætningsforled fra komposita-filen og ordbøger/glossarer, der behandler det pågældende ord, fra glossarfilen. Hermed er artikelen klar til at forlade førsteredaktørens bord.

### Redaktionelt samarbejde

Databasen byder på særlige muligheder, når flere redaktører vil forsøge sig med redigering af det samme intakte citatmateriale. Når førsteredaktøren af en artikel har lavet sit udkast, gennemgås det grundigt af andenredaktøren. I databasen kan andenredaktøren komme med egne forslag til beskrivelse af ordet og til strukturering og anden tilskæring af artikelen, og han kan med få tryk på tastaturet ved hjælp af egne indekseringsfelter afprøve sine ideer. Oftest, i de mange mindre artikler, viser han dog sine forslag blot ved at gøre antegninger i en arbejdsudskrift.

Arbejdsudskriften kan bestå ene og alene af artikelens skelet, dvs. dens afdelinger og definitioner; eller den kan have den længst mulige form, vise alle citater, også dem, der foreslås bragt som nøgne henvisninger, og dem, der foreslås udeladt, samt alle interne kommentarer; eller udskriften kan være udformet præcis som en artikel i den trykte bog.

Samme mulighed som andenredaktøren har den engelske redaktør, der gennemser alle artikler, først og fremmest med henblik på de danske/engelske definitioner. Den engelske redaktør kan vælge at arbejde i et særligt skærmformat, hvor engelsk tekst praktisk kan indsættes og kommentarer tilføjes. Undertiden gives der danske/engelske kommentarer eller oversættelser nede i selve citaterne, også til det formål er der udviklet skærmformater, og et lille program finder vej til de få steder, det drejer sig om.

### Fra database til trykt bog

Takket være databasesystemet kan ordbogens medarbejdere indsamle og indskrive materiale, automatisk indhente oplysninger fra tilkoblede registre, analysere på materialet, redigere det, samle baggrundoplysninger, diskutere med hinanden, gøre interne notater, alt sammen på en særdeles praktisk måde.

Arbejdet foregår inden for et par simple business-databasesystemer: dBASE III og SmartWareII. Vi vil, på vej til fotosætningen, benytte tekstformatteringssystemet TEX, som også blev benyttet ved fremstillingen af registerbindet (interface: Postscript). Helt endelig beslutning om typografi behøver man ikke træffe før sent i forløbet. Hvis bare man har stoffet fornuftigt struktureret og fordelt på diskrete kategorier, kan alt lade sig gøre typografisk. Tegnproblemer og alfabetiseringsproblemer er løst. Systemet er helt åbent. Det kan endnu ikke siges at fungere smidigt, men det vil blive smidigere efterhånden, og snart har vi, hvis vi skulle ønske det, et godt grundlag for at kunne definere et evt. skræddersyet ordbogssystem. Det, der ligger fast, er, at vi hele tiden grundigt kender og kan overskue og håndtere både system og materiale, og at vi hele tiden arbejder henimod nøjagtig det samme produkt. Som nævnt kan vi trække alskens arbejdsudskrifter ud af databasen undervejs, og vi kan definere et hvilket som helst skærmformat, men vigtigst af alt kan vi til sidst direkte ud af én og samme base trække det destillerede produkt: ordbogsartikelen og dermed ordbogen.

Som vi fik det afprøvet under arbejdet med registerbindet, hvor vi gik så direkte som overhovedet muligt fra database til trykt bog, indebærer en sådan bogproduktion indlysende praktiske fordele: ingen ventetider og mellemstationer, ingen traditionel tekstbehandling (det vil fx. sige ingen tegnsætning mellem de forskellige slags oplysninger) og ingen korrekturlæsning i traditionel forstand. Redaktørerne kan i stedet og indtil sidste øjeblik koncentrere sig om den vigtigste indholdskorrektur i selve databasen.

Vi har bevidst udskudt at arbejde med de allermost komplicerede artikeltyper, det gælder verberne med deres partikler og konstruktionseksempler. Men systemet er gjort klar til modtagelse og manipulering af også komplekse strukturer, og ikke mindst, når vi skal til



at håndtere disse mere tunge sager, har vi gode forhåbninger til databasesystemet både som opbevaringskasse og som redigeringsværktøj.

Når vi bliver spurgt, hvad vi vil stille op med de helt store ord, fx. de store præpositioner på over 1000 sedler, om vores databasesystem virkelig kan håndtere dem, så må vores svar være: I krig og kærlighed gælder alle kneb, og stillet over for særlig store leksikografiske problemer bør overhovedet ingen metode foragtes. Problemerne skal søges løst og artikelen præsenteres på den bedst mulige måde. Men grundtanken må fastholdes. Ikke nok med at databasesystemet er et fint stykke værktøj for mangen en videnskabsmand. En ordbog er en database, og i bund og grund betaler det sig at lade den være det fra først til sidst.

### Litteratur

- Nordisk sprogsekretariats rapporter 14. 1991. *Lexikonord*. Oslo  
Ordbog over det norrøne prosasprog. 1983. *Prøvehæfte*. København  
Ordbog over det norrøne prosasprog. 1989. *Registre*. København  
The Arnamagnæan Institute & Dictionary. 1990. *Bulletin 17*. København

## Fra seddelsamling til database: Leksikografisk analyse af islandske verber

Ved *Orðabók Háskólans* — Leksikografisk institut — i Reykjavík arbejder man på en stor historisk ordbog over islandsk. Hovedkilden er en stor seddelsamling der er baseret på traditionel excerpering af tekster samt andre mindre særsamlinger. Disse samlinger bliver siden kompletteret med materiale fra maskinlæsbare tekster når det viser sig at være brug for flere og/eller bedre eksempler for at få frem en fuldstændigere beskrivelse af ordenes mening og brug. Som første trin i redaktionen bliver der lagt hovedvægt på verbernes analyse og udarbejdelse. I det følgende beskriver vi analyseprocessen og redaktionsarbejdet og fortæller om vores erfaring af at forene ny teknologi og arbejdsmetoder med gammelt materiale.

### 1 Indledning

Efterhånden som databehandling vinder indpas i forbindelse med udgivelse af ordbøger, mister de traditionelle seddelsamlinger deres værdi og bliver kun vidnesbyrd om en gammel arbejdsmetode. De håndskrevne og ofte næsten ulæselige sedler afløses af databanker. I mange lande har man historiske ordbøger, og selv om de trænger til revision, nyder man godt af at have en gennemarbejdet ordbogstekst som grundlag.

Da databehandlingen holdt sit indtog var man i Island endnu ved at indsamle materiale til en stor historisk ordbog over islandsk, overvejende med den traditionelle excerpering af skrevne kilder. Redaktionsarbejdet var ikke påbegyndt, og der forelå ingen endelige planer om udformningen af den ordbog, hvis udgivelse man arbejdede hen imod. I det følgende vil vi prøve at gøre rede for, hvordan vi i Leksikografisk institut har forsøgt at forene nye metoder med gammelt materiale i det forberedende arbejde til ordbogen.

### 2 Leksikografisk instituts seddelsamlinger

Da man i 1983-84 tog datateknikken i brug på Leksikografisk institut, besluttede man at gøre et lille ophold og forsøge at få et overblik over det indsamlede materiale og undersøge mulighederne for at gå i gang med redaktionsarbejdet. Man havde indsamlet materiale i næsten 40 år, og resultatet af indsamlingen var nogle kæmpestore alfabetisk ordnede seddelsamlinger:

*Skriftsprogsamling:*

2.2 millioner sedler med eksempler på ord og ordbrug i skriftsprog fra 1540 og til i dag som er blevet samlet med traditionel excerpering, hovedsagelig af trykte bøger.

*Talesprogsamling:*

Ca. 190 tusinde sedler med kommentarer til ord og ordbrug i talesprog der er blevet samlet gennem forespørgsler, især i radioen.

*Andre mindre samlinger:*

Sedler fra utrykte ordsamlinger fra det 18., 19. og 20. århundrede.

Som forholdene var, forekom det indlysende at man ved indførelsen af edb i første omgang burde koncentrere sig om selve seddelsamlingerne, snarere end redaktionen af ordbogsteksten som sådan. Med dataregistrering af samlingerne sigtede man efter at skaffe sig viden om deres omfang og struktur, inden man gik i gang med selve ordbogen for derved at få en bedre bearbejdet tekst. Man besluttede at gå i gang med to opgaver: For det første udarbejdelsen af et fuldstændigt dataregister over ordforrådet i skriftsprogsamlingen, og for det andet en nærmere bearbejdelse af visse afgrænsede dele af hele samlingen.

Man gik først i gang med udarbejdelsen af førnævnte oversigtsregister. Foruden selve ordene har man også registreret fundamentale oplysninger om dem. Dette arbejde er nu færdigt, og det har bragt forskelligt for en dag vedrørende ordforrådet og samlingens karakteristiske træk. Selv om registeret kun giver begrænsede oplysninger, kaster det lys over visse punkter, som tidligere tiders ordbogsfolk ikke havde nogen mulighed for at gøre sig bekendt med og giver adgang til samlingen fra betydeligt flere sider en før. Vi har nu, bl. a., en ganske klar idé om skriftsprogsamlingens omfang m.h.t. ordforråd og dets inddeling i ordklasser (se figur 1 og 2).

### Skriftsprogsamlingen (1540-1991)

Antal eksempler	2.200.000
Antal ord	610.000
Eksempler pr. ord:	
1	62%
2-4	23%
≥ 5	12%

Figur 1: Skriftsprogsamlingens omfang.

Ordklasse	Verbum	Substantiv	Adjektiv	Adverbium
Antal ord	16.000	500.000	84.000	6.000
% af ordforrådet	3%	82%	14%	1%
1 eksempel pr. ord	44%	64%	56%	50%
2-4 eksempler	22%	23%	25%	26%
≥ 5 eksempler	34%	13%	19%	24%
Grundord	35%	5%	12%	35%
Præfiks/sammensat	65%	95%	88%	65%

Figur 2: Skriftsprogsamlingens sammensætning.

Med hensyn til opgave nr. to, nemlig en nærmere bearbejdelse af en afgrænset del af samlingen, valgte man verberne. Man gik i gang med at dataregistere selve eksemplerne, samt at analysere dem med henblik på udarbejdelsen af en ordbogstekst. Dette arbejde er endnu langt fra færdigt, men ikke desto mindre har man allerede nu en langt bedre oversigt over samlingens karakteristiske træk, end man fik gennem oversigtsregisteret. Hvis man i korte træk vil gøre rede for samlingens almindeligste karakteristiske træk, kan de siges at være følgende:

- \* Et stort og mangfoldigt ordforråd.
- \* Righoldige vidnesbyrd om orddannelse.
- \* Mangfoldighed m.h.t. betydning, bøjning, syntaks og ordforbindelser.
- \* Eksempler der er værdifulde langt ud over det leksikalske (fx forskellige kulturhistoriske udsagn).

De her nævnte træk er på samme tid samlingens stærke og svage side og indebærer visse mangler, som man må tage stilling til. Inden vi kommer nærmere ind på dette, er det dog nødvendigt at sige lidt om, hvordan man har grebet analysen an og om forbindelsen mellem denne og en påtænkt ordbogstekst.

### 3 Bearbejdelse af samlingerne

#### 3.1 Analyse af materiale

Ved bearbejdelsen af instituttets materiale har man hidtil udelukkende koncentreret sig om verberne, og ved udarbejdelsen af analyseprogrammer har man taget sigte efter dem. Analysen skal for det første være tjenlig i forbindelse med udarbejdelsen af ordbogstekster, og for det andet skal den være så nøjagtig, som det er muligt. Databasen skal senere være tilgængelig for såvel forskere som publikum. Det er derfor ønskeligt, at oplysningerne om ordforrådet er så udførlige som muligt, selv om de ikke alle kommer frem i ordbogen. Da analysen foretages i etaper, kan visse dele af den også komme til nytte ved redaktion, selv om de ikke

forekommer i den endelige tekst. Det er praktisk at kunne inddele eksemplerne efter forskellige kendetegn, der allerede er blevet analyseret for at gøre det nemmere og hurtigere at analysere på senere trin.

Vi vil ikke her komme ind på analyseprogrammerne som sådanne. Man arbejder i øjeblikket med ændringer af disse i forbindelse med en ny database, der skal tages i brug, og samtidig foretager man en omvurdering af selve analysen og dens tilrettelæggelse. Analysen forbliver dog i princippet uændret og kan derfor beskrives her i store træk.

### 3.2 Verbalanalyse

Det er meningen at eksemplerne på ordenes brug skal udgøre den vigtigste del af ordbogsteksten. Bearbejdelsen går derfor først og fremmest ud på analysen af eksemplerne. Hvert enkelt eksempel analyseres systematisk, og analysen af verber kan inddeles i fem hovedpunkter:

- \* *Eksemplets aldersbestemmelse.* Den kan udføres maskinelt efter kildefortegnelse.
- \* *Morfologisk analyse af verbets form i det foreliggende eksempel* (modus; form; person og tal; tempus).
- \* *Syntaktisk analyse af verbets tekstuelle omgivelser.* De sætningsled, der på en eller anden måde knyttes direkte til verbet, analyseres m.h.t. form og betydning. Der er her tale om subjekt, objekt — et eller to, samt adverbier og præpositionsled der er nært knyttet til verbet som en slags fast bestemmelsesled (se figur 3).
- \* *Semantisk analyse af verbum eller verbumsforbindelse.*
- \* *Andre punkter i nødvendigt omfang:*
  - faste ordforbindelser
  - ordsprog
  - kulturhistorisk beskrivelse, forklaring vedrørende ordbrug m.m.

Efter analysen sørger datamaskinen for inddeling af eksemplerne og ordner dem efter visse regler, hvorefter et ombrydningsprogram, der er tilkoblet databasen sætter ordbogsteksten i overensstemmelse med analysen. Figur 4, hvor man kan se en prøvetekst af verbet *péna* (tjene), viser resultatet. De syntaktisk analyserede afsnit "oversættes" maskinelt som hoved til vedkommende artikel, - d.v.s. at datamaskinen selv udformer redaktionel tekst, fx *péna e-m til e-rs, péna e-u* o.s.v. Tekstens udformning bestemmes især af verbets form og dets tekstuelle omgivelser. Beskrivelsens hovedafsnit tager sigte efter visse morfologiske og syntaktiske træk: aktiv, medium, (selvstændige) participier; argumenter (tilstedeværelse og type af subjekt og objekt(er)). Disse hovedafsnit kan så blive delt i mindre afsnit m.h.t. andre led der er nært knyttet til verbet (ordnet alfabetisk efter præpositioner og adverbier).

## Syntaktisk og morfologisk analyse

A. eptir sem uppfynding þessi þjenar fólkinu til virðingar og þrýðis.	B. Á ég engar skræður, sem þéna þessu guðspjalli?
Kilde: JólGrvKlím. 157 Alder: 18m	Kilde: JThSk. II, 52 Alder: 19m



Verbum:

	Form	Modus	Tempus	Tal	Person
A. þjenar	aktiv	indikativ	presens	ental	3
B. þéna	aktiv	indikativ	presens	flertal	3

Subjekt:

	Kasus	Betydning
A. uppfynding þessi	nominativ	-levende AKT
B. [skræður]	nominativ	-levende BOG

Objekt:

	Kasus	Betydning
A. fólkinu	dativ	+levende MENNESKE
B. þessu guðspjalli	dativ	-levende TEKST

Adverbial:

	Type	Ord	Nominal
A. til virðingar og þrýðis	præp.led	til	genitiv/-levende
B. —	---	---	---




---

A. þéna e-m til e-rs  
B. þéna e-u

---

Figur 3: To eksempler på verbet *þéna* og hovedtrækkene i den morfologiske og syntaktiske analyse af dem.

Betydningen er underordnet i forhold til strukturen. Hver artikel får sin betydningsbeskrivelse, og hvor der er tale om mere end én betydning, indeles vedkommende artikel efter nummererede betydninger. Ordforbindelser placeres i artikler, hvor de anses for at høre hjemme m.h.t. struktur og betydning men får (som oftest) en særlig betydningsbeskrivelse. I hvert enkelt afsnit eller artikel ordnes eksemplerne efter alder således at de ældste eksempler kommer først og de yngste sidst (det er angivet foran hvert eksempel fra hvilket århundrede og hvilken trediedel af århundredet det stammer, således at et eksempel fra 1540 kendetegnes ved *16m* og et eksempel fra 1921 kendetegnes ved *20f*).

I verbets hoved gives oplysninger om bøjning foruden andre punkter, der vedkommer verbet i det hele taget.

### 3.3 Vurdering af materialet

Først efter analysen kan man danne sig et billede af det materiale der er indeholdt i seddelsamlingen og vurdere beskrivelsen, bl.a. i sammenligning med foreliggende kilder om det islandske sprog, især ordbøger. Man må tænke på, at edb-analysen gør det muligt at fremstille en tekst, der er mere systematisk og udførligere analyseret, end man tidligere havde mulighed for at gøre. Samtidig kommer der visse skævheder, lakuner og slagsider til syne, som skjuler sig nemmere i en mindre systematisk og mere kompakt tekst. Det kan dreje sig om:

*Tvivilsspørgsmål* der skyldes at de foreliggende eksempler ikke er klare nok. Mange af eksemplerne i seddelsamlingen er så knappe, at de ikke forslår til at vise syntaks eller betydning klart og utvetydigt. Eksempler (ikke mindst fra poesi) kan også være vanskelige at forstå eller måske helt uforståelige. Endvidere er nogle eksempler tvetydige m.h.t. syntaks og/eller betydning.

*Lakuner* i beskrivelsen, hvor der mangler eksempler på strukturer og/eller betydninger som man ved om. Det kan enten være sådan at vi slet ingen eksempler har eller at vi ikke har eksempler fra et bestemt tidsrum.

*Skævheder* af forskellig slags, der ser ud til at findes i beskrivelsen. Antallet af eksempler på enkelte strukturer og/eller betydningsvarianter kan se ud til at give et forvansket billede af deres stilling i sproget. For det første kan der være forholdsvis mange eksempler på det, der må anses for at være forholdsvis sjældent og få eksempler på mere almindelige strukturer og/eller betydninger. Eksempelernes aldersdistribution kan også virke unaturlig i forhold til, hvad man kunne forvente. Og til sidst kan forekomsten af bøjningsvarianter se ud til ikke at give det rigtige billede m.h.t. alder, hyppighed o.s.v., — oftest derved at der er flere eksempler på forskellige afvigelser, end man kan anse for at være naturligt.

Sådanne skævheder må så vidt muligt rettes, selv om ordbogstekstens fremstilling indebærer en vis fare for slagside.

Det er meget forskelligt fra det ene verbum til det andet, hvor fyldestgørende en beskrivelse seddelsamlingen indeholder. Det er også forskelligt, om der helt klart er tale om lakuner og/eller skævheder, eller om der blot er tale om mistanke om sådanne. Det ville være ønskeligt at efterprøve seddelsamlingens vidnesbyrd for at stadfæste det billede, som det

giver, eller for efter omstændighederne at gøre det mere fuldstændigt. Datateknikken har gjort dette muligt, i hvert fald indenfor visse grænser, med søgning i maskinlæsbare tekster. Vi vil til sidst komme lidt ind på vores erfaring m.h.t. en sådan søgning efter verber i den tekstsamling, der forefindes på instituttet.

## 4 Tilføjelse af materiale

### 4.1 Tekstsamlingen

Leksikografisk institut har en ganske stor tekstsamling, der ikke desto mindre er begrænset i mange henseender. Den består nu af 171 tekster med ialt godt og vel 15 millioner løbende ord (ca. 80 Mb). De fleste tekster er ganske nye (fra de sidste 6-10 år) men samlingen indeholder dog også én tekst fra det 16. århundrede, to fra det 18. og én fra det 19. århundrede, foruden nogle gamle tekstfragmenter. Teksterne er endvidere ensidige m.h.t. struktur og indhold, langt den største del af teksterne stammer fra nogle få emnekatogier (romaner, biografier eller erindringer men også Bibelen og den islandske lovsamling i sin helhed).

Det har kostet forholdsvis lidt besvær og få udgifter at tilvejebringe tekstsamlingen. De fleste tekster har man fået i maskinlæsbar form fra bogforlag og trykkerier, og det forklarer til en vis grad ensidigheden. Den søgning efter eksempler, som vi har foretaget indtil nu, har til dels været forsøgsvirksomhed, og man har planer om at tilvejebringe en mere systematisk og mere varieret samling, når man har høstet lidt flere erfaringer.

### 4.2 Søgning i teksterne

På Leksikografisk institut anvendes operativsystemet Unix og tekstsøgeprogrammet GNU-grep fra Free Software Foundation. Det søger bogstavrækker i tekstsamlingen og giver linien med det ord som man søger foruden de to forudgående og de to følgende linier. Man må søge alle ordets bøjningsformer, og af dem kan der være temmelig mange i islandsk. Man laver en liste over alle bøjningsformer, der forekommer i hvert ord og søgeprogrammet søger siden i overensstemmelse med listen. Dette er forholdsvis enkelt, selv om det er ret forskelligt hvor mange rækker der skal opgives. For et regelmæssigt verbum med svag bøjning må man søge efter 30 forskellige bogstavrækker, men varianter i ortografi og/eller bøjning kan forøge antallet betydeligt, — helt op til godt og vel 100 former.

Når der findes mange eksempler, er det et forholdsvis stort arbejde at gennemgå hele samlingen af disse, sortere dem og udtage eksempler på andre ord, end det man søger. Der er en hel del homografiske former, og i de fleste tilfælde kommer der en del uvedkommende eksempler frem ved søgningen. Det er nødvendigt at være på vagt over for dette, for nogle homografiske former er temmelig besværlige at søge af denne grund. Skal man fx søge verbet *eiga* (have), må man passe godt på, idet én af verbets former er *á*, og er derved homonym med en af de hyppigst forekommende præpositioner i sproget, foruden at *á* også optræder som visse former i to forskellige substantiver. I en lille del af tekstsamlingen viste der sig at være godt og vel 12.500 eksempler på *á*, og heraf var kun knapt 3% verbalformen *á*. Til trods for enkelte eksempler af denne art sker det forbavsende sjældent, at homografiske former volder besvær.



Resultatet af søgning er meget forskellig, efter hvilket verbum der er tale om, og tekstsamlingen viser sig at være nyttigst ved analysen af verber i gennemsnitsstørrelse. Når verbet er meget hyppigt forekommende koster det et enormt arbejde at gennemgå eksemplerne og resultatet svarer absolut ikke til anstrengelserne (i en lille del af samlingen var der fx godt 22.000 eksempler på verbet *vera* (være)) og hvis verbet er forholdsvis sjældent kan det være tvivlsomt, om det overhovedet findes i tekstsamlingen. Resultaterne af tekstsøgning kan være tilfredsstillende af to forskellige grunde:

- \* selv om den ikke giver noget nyt, kan den være nyttig fordi den giver bedre eksempler end dem, man havde i forvejen eller en mere fyldestgørende aldersdistribution end den, der findes i seddelsamlingen;
- \* den giver rene tilføjelser til seddelsamlingen, dvs. eksempler som viser sætningskonstruktioner, betydninger eller ordforbindelser som ingen eksempler fandtes på i forvejen.

Vi skal nu se på to konkrete eksempler på resultater af søgning. I første tilfælde kom der intet nyt frem, selv om søgningen ikke var helt uden gavn. I sidste tilfælde havde seddelsamlingen øjensynligt givet en meget mangelfuld beskrivelse, og her gav søgningen gode resultater.

#### 4.2.1 *péna*

Verbet *péna* (tjene) er forholdsvis sjældent i islandsk nutidssprog. Ikke desto mindre fandt man ialt 51 eksempler ved søgning i tekstsamlingen. Selv om disse ikke viste nogen strukturer eller betydninger, som der ikke var eksempler på i forvejen, fandt man eksempler, der gjorde seddelsamlingens beskrivelse mere fuldstændig m.h.t. aldersdistribution. I 7 af 19 artikler fik man nye eksempler fra tidsrum, som ikke var repræsenteret i seddelsamlingen. Det er bemærkelsesværdigt, at de færreste af de nye eksempler er fra det 20. århundrede, selv om de fleste tekster er fra dette århundrede, og de fleste af de ældre eksempler kommer fra tekster, der allerede var blevet excerperet. Endvidere kom der nye eksempler, som måtte anses for at gøre ordbogsbeskrivelsen bedre, uden at de direkte udfyldte lakuner. Af de 51 eksempler, der blev fundet i tekstsamlingen, er 8 benyttet i verbets beskrivelse (d.v.s. i prøveteksten i figur 4).

**béna** -tí/-að (nt. alltaf -ar), -t/-að  
Sjá Orðabók Fritzens.

**ÞÉNA** = 1. *bjóna* [e-u, e-m] 16m og hefði gjarnan kent þeim huern þat Laugen ætti at þjena. CorvPost., II 11 18s "hvær þénti þýlgjarr, / enn þú? Satan! MiltPar., 122 19f Sá kann ekki að stjórna, sem ekki hefir lært að þéna. CJ., 286 19f Sælir eru þeir sem vel hafa þént (af hafa lokið). CJ., 332 #2. *vinna, starfa, gegna starfi* d áku, stað e. í áku *embætti, vera í vísit e-rs staðar* 18f þann tíma Jón Magnússon þénti í því gestlega standi prestsemþættinu. Alþb. IX, 524 (1709) 18s 19f hann má ei þéna í réttinum nema í forðillum bróður síns. Fjalik. 1885, 79 (um 1800) 19m [hann] þénti í þeim stórnum til dauðadags. CJ80ist., 49 (1863) 20m Hún hafði ... þénað í húsum hjá fyrirmyndarfolki. VSVFimm., 52 20m Hjá honum hafði Magnús þénað og numið um 6 ára skeið. Arsrtsf. 1959, 72 #3. *nýast, koma að gagni; eiga við* 18m uppi Framkyrkjuni eru alls 16 bord sem þjena fyrir Loftf. Þps. Áll 21, 6 (1758) 16m Til er eirnenn 1¼ af raufþryktu Cartune sem þena kann í alltaresduk. Þps. Áll 21, 214 (1761) 19m sú aðkomandi stúlka fór höndum um konna er á gólfu lá eftir því er þénaði og hún vissi að með þurft. JÁBj. 1, 17 20 mætti sumt í þessum ritdóm! þéna sem spegill, að vísu mjög skjáður, er aðrir rithöfundar gæta séð í brot af sjálfum sér. ÞórbþEin., 175 20m [glæraugun] hafa þénað betur en glæraugun mín, og nú er ekki til neins að leita. ÞórbþEið., II, 342 20s En „þó hér séu litlar framfarir þá hef ég þó séð ýmislegt er ég hefði ekki áður séð. Hefur mér þá alltaf orðið að spyrja hvort þetta gæti ekki þénað heima!“ SvipAthn., 183 #4. *afila e. græða peninga* 20m ég vona að þú þénir vel í vetur, svo að þú getir sent mér aura. ÞjórnMát., 251 20s Þegar ég hefði lokið verkefni mínu í nóvemberlok hafði ég þénað mjög vel. JóhSvEisk., 207 20s andskotinn og einstaklingurinn eiga að þéna en alþýðuræfilinn fær ekki neitt! LÁbergBlatt., 188

málsh.  
málsh.

þéna af e-u: *valda e-u, stuðla að e-u* 19s Það sem þjenar að hreifingunni [þ.e. í vatnsmyllu], Nf. XI, 60  
þéna á e-u: *græða (f) á e-u* 20m menn eru nú vanir að gefa fremur með sér, heldur en þéna á því. ÞorEinkinn., 89  
þéna til e-rs: *koma að gagni e. nýfast við e-ð; stuðla að e-u* 16m með þú / þat í holddino at lífa þjenar meir til a usgatar at af þa veit eg eigi huort eg skal kiosa. Fil. I, 22 (OG) 16m Enn andi er sa hinn innre madur og yre / sem líufer og verkar þad sem til andigra hluta og epter komandi lífdaga þjenar. Rómform., a II # (OC) 17f að því Leite sem [vonin] þjenar til Syndnanna fyrrengfingur. VidPost. II, 303f 18m í þessu lande var yferfjórstanligt af öllu því sem þjenaði til mannsins gagns og hæginda. JóhGrvKlim., 188 18ms það þénar, og til að gjöra greinarmun milliun dyggða og ódyggða. ÞPStJÁBj., 4 19s "að þér streymj jafnan / all, hvað þénar yndis til. GPHSKáttk., II, 170 (19. öld) 19m Sér hún þá til hvers [hrifurinn] þénar og tekur feginlega á móti honum. JÁBj. 1 III, 261 19m Flest vilji þið mér fyrirnuma sem til skemmtunar þéna mætti. JÁBj. V, 90 19s gaf [hann] honum fullnagt til að taka sér fyrir hendur það hann ákisti þéna til að málinu yrði framgengt. FréggFyig., I, 313 20 efturfarandi ráðleggingar munu þó geta þénað til að lína nokkuð þrautir sjúklingsins. ÞórbþEin., 9 20m Enginn vissi til hvers þetta næstum útmáða mannvirkí hafði þénað í lífinu. ÞórbþStein., 224 20m allt sem þénar til upp lýsingar um síða og kristni hald er dyggð. Þjórnjómf., 61 20m Ég hef hugsað mér að bjóða dönsku stjórnminni mína uppfinningu ... en ég er þó hræddur um, að það þéni ekki til neins. JónHelgMann., II, 138

A+

A+

A+

A+

væri á smávörðinni. Austri. 1887, 13 20m Það eru meiri lífandis ósköpin, sem allir geta þénað á þessu ströð. StSteinGlugg., 27 20s Það var sumaríð '52, ég ... ætlaði að þéna einhver lífandis ósköp á því að salta síð. VéstLÉfir., 22 20s Mennirnir rveir ... státu hæglafrí við óðýran drykk ... Það væri ekkert hægt að þéna á þeim. 8ónapart., 161

**ÞÉNA E-U**: #1. *bjóna e-u, styðja e-ð, eiga við e-ð e. falla að e-u* 17 "þó vil eg gjarnan þéna hvers manns sinni / vísit í góðu, en vondur ei. JSamsKD. II, 52 (17. öld) 17s Noturnar og Tomarnar eiga að þjena Orðunum. Grad. A, III / 18m stundabe mitt eggi gagn und þeim tilslætte, at eg vilde þar með þjena almúgans tilstande. JóhGrvKlim., 136 19m Á ég engar skráður, sem þéna þessu guðspjallfríthsk. II, 156 20m toppfallsblakir, svo og allt fast í reidönum, sem daglega þénar lausum reidönum, er tekið niður ... vor og haust. Verksjóv., 101 20m lengi hafði verið svipazt um eftir drottningarefni, sem þénað gat tilgengi frúarinnar. ÞerjMann., 311 #2. *starfa e-rs staðar; þjóna e-u (með starfi að e-u embætti e. viðfangsefni)* 18f að lögmaðurinn Lauritz Gottrup dragi sig aldeilis undan að þéna réttinum í morðsök þeirri, sem víðvikur Magnósi Benediktssyni. Alþb. IX, 378 (1705) 18f [svo] presturinn, sem þéni Kolbeinsstaðakirkju [mætti ívein nóttum hvílast, þá hann fremur þar þjónustu vegna fólkins nauðsynja]. JHBSk. I, 132 #3. *koma að gagni e-rs staðar, nýfast e-u* 18f Reifingarrista þjargleg, brúkast so sem heimilinu þjenar. Járðab. XI, 27

**ÞÉNA E-M**: #1. *þjóna e-m, koma e-m að gagni* 17m "Loddurum þéni listin sú, / að láta folana gesa. JGubnÁrm., I, 43 17s "all flestar frúr um Danmerki ridi / voru sóttar at henni þéni. lík. II, 40 18 "Læzt ég vanda legstaðni, / — lengi klárin þéni mér. HérabSgf. II, 276 (18. öld) 18m Þeir voru eingvum lögum underopner, heldur þjenaði þeim sé eintóme keisarans vilje fyrer lög. JóhGrvKlim., 189 19f með hvörju megum við hafa þá eru að þéna Ydur? MSIGÁh. I, 84 19f Sá, sem ofmjög þénar öðrum, fortaerir sjálfum sér. CJ., 296 19f Sá þénar öðrum nóg, sem gjörir það af alvöru. CJ., 302 20 til hvers er verið að þessu? Hverjum þénar það? ÞórbþEin., 158 20m Ein stór og mikil líteglu, vandlega girt og vel sterk, þénaði öllum vinnuöðnumum. GHagaVrk., 23 #2. *hugfa e-m, koma e-m vel* 18m so hefði [minn lereimastari] að vísu kendt mjer þat sem þjenaði einum áhvöndi kyrkjuþjenara. JóhGrvKlim., 234 19m að enginn býði sér ... neins konar hægingdi sem ferðamanni þénaði. JÁBj. II, 227 19s Þú erst svo rjóðsettur og gæðinn og sér strax allt hvað þóðum þénar. JGubnSKth. I, 34 (1879) e-m er þént með e-ð: Tms Hér er nú ekki þént með það, ég er til ekkert hrifin af þessu, langar ekki til einhvers. S-Möl. e-m er þént með e-u: *e-ð kemur e-m vel, e-ð hafir e-m* 17m "Samt er pier ekki þént með þú, / þo eg þig aukska villdi. HPPass. XXXI, 12 18m Hvad er okkur þa þient með því að binda öðrum í milli Einka-Maal. ÞKetGL. 36 19f Vænjanlega losnar hér okkar embætti, sem þér er þént með. JHúst., 70 (1819) 19fm vore ydur þient með hafu Gemulufalle ... þa skyldde eg þera mig að falast eftir því. SafnF. XII, 223 19m Eg verð enn að skorast undan verki þessu, því eg treysti mér alls ekki til að gjöra það svo, að þinginu verði þént með því. Alþ. 1855, 794 19m hvermin kirkjum góða prestum gesti nú leingur verið orðið þént með eður hagkvæmara, að fá greiðslur sínar í ull og tóg. Alþ. 1849, 361 e-m er þént með e-m: *e-r gagnast e-m* 19s mér er þetta bending um, að stjórnmíni sé ekki þént með mér. ÚrtjÁ. II, 199 (1873)

málsh.  
málsh.

þéna e-m til e-rs: *koma e-m að gagni við e-ð; stuðla að e-u* í þágu e-s 17m "hvad þier til frelses þjena kann: / þad er fullkonnad seigir hann. HPPass. XLII, 14 18m eptir sem uppáfynding þessi þjenar fólkinu til virðingar og þryðis. JóhGrvKlim., 157 19m ef „filgjandi“ sögukorn „mætti þéna“ iður til skjemtanar. Fjölín. II, 2, 57 19m Ég meina að Edda miður eigi að þéna skáldinu til létts en til að þryða hér og hvar verk hans. Rætur., 235 (1835) 19m því allir hlutir þjena þeim til góðs, sem elska guð. MTTæfif., 97 19s hagleiddi jafnframt, hvernig hin líflimótelegustu arvik geta þénað mönnum til góðs. FréggFyig., I, 117  
þéna sér e-ð: *vinna sér e-ð inn* 18m að dimissi frá skólanum mættu ekki þjena sér braud hjá verðslögum mönnum skikkantlegum. ÞorKÆt. I, 26 (1740)

A+

A+

**ÞÉNAST**: e-m þénast e-að: *e-r græðir e-ð* 20fm svo mikið þénað aðist honum af alls konar sveitavöru á flakki sínu. RaubSk. IV, 92

**ÞÉNANDI**: sem kemur (e-m) að gagni 19f [hún er] samilega kunnandi til vellestra almúga þénandi og nauðsynlegra kvenmannsverka. JónHelgMann. III, 91 (1818)

Figur 4: En þröng þvætten dær viser verbet þéna sådan som man tænker sig at teksten i den historiske ordbog kommer til at se ud.

#### 4.2.2 *ljúka*

Det andet eksempel på søgning i tekstsamlingen angår verbet *ljúka* (blive færdig med, slutte). Det var helt klart, at seddelsamlingens beskrivelse var langt fra at være fyldestgørende, selv om der var 126 eksempler, bl.a. manglede der eksempler på almindelige sætningskonstruktioner.

Verbet er forholdsvis almindeligt i sproget, og ved søgningen fandt man knapt 4.700 eksempler af forskellig art. Figur 5 viser eksemplernes inddeling m.h.t. sætningskonstruktion sammenlignet med eksempler fra seddelsamlingen. Ved søgning i tekstsamlingen fandt man 41 sætningskonstruktioner — heraf 21 der ikke forefandt i seddelsamlingen. Seddelsamlingen har derimod eksempler på 6 strukturer, der ikke kom frem ved søgning. Nogle af de strukturer der kom fra tekstsamlingen, fandtes i gamle tekster, som allerede var blevet excerperet, uden at de blev taget med. Eksempler på *ljúka við e-ð*, der er en meget almindelig sætningskonstruktion, fandtes bl.a. i Det nye Testamente fra 1540.

Foruden de nye strukturer gav søgningen mange ordforbindelser, der ikke fandtes i seddelsamlingen, og de eksempler der fandtes ved søgningen viser også betydeligt flere betydningsvarianter end man havde i forvejen.

Disse resultater viser, hvilket gavn man kan have af at bruge tekstsamlingen sammen med seddelsamlingen, idet man i hver kilde for sig finder materiale, der ikke findes i den anden.

### 5 Konklusion

Til sidst vil vi sammenfatte nogle hovedpunkter og gøre opmærksom på, hvilken lærdom man kan drage af dem.

- \* Leksikografisk instituts seddelsamlinger indeholder rigt varierede eksempler på brugen af et stort antal ord, og her findes materiale, som det er vanskeligt at komme til på en hurtig måde.
- \* Edb-analysen gør det muligt for os at fremstille en udførligt analyseret og systematisk ordbogstekst, der giver et godt billede af seddelsamlingens karakteristiske træk.
- \* Forskellige punkter må underbygges med flere kilder, og der mangler især eksempler på herskende sprogbrug.
- \* Tekstsamlingen har til trods for sin ufuldkommenhed vist sig at være gavnlig til at udfylde lakuner og rette de beskrivelser, der findes i seddelsamlingen, især når det drejer sig om de mere almindeligt forekommende fænomener.
- \* For at have virkeligt gavn af søgning i tekstsamlingen uden at arbejdet bliver alt for overvældende er det vigtigt, at man gør sig klart, hvilke fordele og ulemper og begrænsninger både metoden og selve samlingen har.
- \* Skal man bruge en tekstsamling målbevidst ved redaktionen af ordbogen, sådan som vi her har beskrevet, må samlingen være fornuftigt sammensat:

- den må ikke være for stor;
- teksterne må spænde over hele det tidsrum, der tages med i ordbogen og være så varierede, som det er muligt m.h.t. emner, stil, m.m.;
- samlingen burde være klassificeret efter forskellige træk, der er karakteristiske for teksten: alder, emne, stil, forfattere (alder, afstamning, samfundsklasse), m.m.;
- det skulle være muligt at søge såvel i hele samlingen som i visse afgrænsede dele af den (dele af teksterne, tekster indenfor visse kategorier, tekster af nord-islandske forfattere fra det 18. århundrede o.s.v.).

- \* Det skal nævnes, at tekstsøgning af den art, som vi har arbejdet med, efter al sandsynlighed kan give bedre resultater, når det drejer sig om verber, end når man har med andre store ordklasser at gøre. Verberne er forholdsvis få (knap 3% af ordforrådet i skriftsprogsamlingen og ca. 7% i en ny frekvensordbog over islandsk nutidssprog (Jørgen Pind et al., 1991)), men de fylder dog forholdsvis meget i tekster (20% af løbende ord i Frekvensordbogen). Man må derfor gå ud fra, at en tekstsamling af en nogenlunde størrelse indeholder eksempler på de fleste verber. Det kunne derimod vise sig at være vanskeligere at finde eksempler på visse andre substantiver end de hyppigst forekommende.
- \* Der er sikkert også brug for flere og mere varierede eksempler med verber end med de fleste andre ord, idet de forskellige sætningsomgivelser er så vigtige i forbindelse med beskrivelsen af deres betydning og brug. Det kan også tænkes, at lakunerne i de andre ordklasser er af en anden art, — at der fx snarere mangler nye ord end flere eksempler med de ord, man har i forvejen — og da er det ikke sikkert, at de her beskrevne metoder kan anvendes.
- \* Konklusionen er, at både gammeldags seddelsamlinger og maskinlæsbare tekstsamlinger har deres svage og stærke sider i forbindelse med den foreliggende opgave og kan derfor supplere hinanden. Samtidig skal det understreges, at seddelsamlingerne er det grundlag, som vi bygger på, og erfaringen har vist os, at de er uvurderlige, under de nuværende omstændigheder.

	<i>Skriftsprogsamling</i>	<i>Tekstsamling</i>
ljúka	1	2
+ af e-u	-	1
+ aftur	-	1
+ um e-ð/e-n	-	2
+ upp	2	32
+ upp fyrir e-u/e-m	-	11
+ úr e-u	2	21
+ út af við e-ð	1	-
+ við e-ð	-	396
ljúka e-ð	2	2
+ aftur	1	1
+ í e-u	-	1
+ um e-ð	1	-
+ upp	3	4
ljúka e-u	15	1.295
+ af	4	36
+ aftur	-	9
+ á e-ð	1	10
+ á e-m	1	-
+ saman	-	1
+ (í) sundur	6	10
+ til	-	1
+ upp	20	270
+ upp fyrir e-m	-	24
+ utan um e-ð	-	1
ljúka e-m e-ð	2	-
+ upp	-	1
ljúka sér	16	37
+ upp	-	2
e-u lýkur	5	2.285
+ á e-u	-	2
+ með e-u	-	50
+ upp	-	4
e-m lýkur	7	31
... lýkur	7	5
... yfir lýkur	-	61
ljúkast	2	4
+ að e-m	1	-
+ aftur	1	5
+ sundur	1	2
+ um e-ð/e-n	-	6
+ upp	6	52
+ upp fyrir e-u/e-m	1	6
+ utan um e-n	-	2
+ yfir e-ð	-	1
+ yfir e-n	-	1
e-u lýkst	1	-
lokinn	6	6
	126	4.697

Figur 5: Antallet eksempler på forskellige sætningskonstruktioner med verbet *ljúka* fra tekstsøgningen, sammenlignet med seddelsamlingen.

**Litteratur**

- Friðrik Magnússon. 1988. Hvað er títt? Tíðnikönnun Orðabókar Háskólans. *Orð og tunga* I:1-49.
- Jón Hilmar Jónsson. 1988. Sagnorðagreining Orðabókar Háskólans. *Orð og tunga* I:123-174.
- Jón Hilmar Jónsson. 1990. A Standardized Dictionary of Icelandic Verbs. I: Jörgen Pind & Eiríkur Rögnvaldsson, red. *Papers from the Seventh Scandinavian Conference of Computational Linguistics, Reykjavík 1989*: 268-296. Institute of Lexicography, Reykjavík.
- Jörgen Pind, Friðrik Magnússon og Stefán Briem. 1991. *Íslensk orðtíðnibók*. Orðabók Háskólans, Reykjavík.

Christian-Emil Ore

## Dokumentasjonsprosjektet ved Det historisk-filosofiske fakultet, Universitetet i Oslo

Det historisk-filosofiske fakultet ved Universitetet i Oslo har det siste året gjennomført et forprosjekt og drevet annet forberedende arbeid i forbindelse med det såkalte dokumentasjonsprosjektet. Dette prosjektet har som mål å bygge opp databaser over arkivmaterialet ved fakultetets samlingsavdelinger. Samlingsavdelingene er oldsaksamlingen (arkeologiske samlinger), folkemusikksamlingen, avdelingene for etnologi og folkloristikk samt leksikografimålføre- og navnegranskingavdelingen ved Institutt for nordistikk og litteraturvitenskap.

Prosjektets idé er å lage databaser for de enkelte fagområdene, men på en slik måte at det er mulig å samkjøre dem. Åpenbare felles innganger til dataene er tid og sted. Gjenstandene i de arkeologiske samlingene har et funnsted og kan som oftest dateres til en periode. Avdeling for navnegransking har arkiv over norske stedsnavn som er sted- og tidfestet. Navn kan også gi verdifull informasjon for arkeologer og historikere, og særlig for språkforskere. Det skulle ikke være nødvendig å drive videre utbrodering av nytten ved muligheten av en krysskobling av arkivene.

Den oppgaven vi har gitt oss i kast med, er ikke av de minste. Samlet er det tale om 12-13 millioner arkivkort og dokumenter, hvorav det meste ligger i de leksikografiske avdelingenes skuffer. Bokmåls- og nynorskavdelingen har hver i overkant av tre millioner ordbokssedler, mens målføre- og gammelnorskavdelingen har omlag to millioner sedler tilsammen. Det meste av arkivmaterialet er uegnet for optisk lesing og må rett og slett skrives av. Registreringsarbeidet vil samlet kreve 600-700 årsverk. Det er klart at det ikke er mulig å få ekstraordinære midler til å ansette registreringspersonell. Tanken har derfor hele tiden vært å bruke arbeidsledige til registrering, og å ansette studenter og ferdige kandidater til å overvåke arbeidet. Denne modellen er nå under utprøving. Sommeren 1991 ble det ansatt fire forskningsassistenter, og i august ble de første 8 arbeidsledige satt i virksomhet.

### Oppbygging av de leksikografiske databasene

For å få erfaring har vi begynt dataregistreringen ved de leksikografiske avdelingene samt ved myntkabinettet. Elektroniske myntkataloger er et interessant felt, men siden temaet er leksikografi, skal vi la myntene hvile og konsentrere oss om ordene. I resten av denne artikkelen vil jeg prøve å forklare hvordan vi har tenkt å løse den formidable oppgaven vi har gitt oss i kast med. Gammelnorskavdelingen og nynorskavdelingen er brukt som eksempler, idet deres materiale representerer hver sin ytterkant av arkivene. Til slutt vil jeg belyse en del problemer og valg forbundet med de datatekniske løsninger. Drøftingen er holdt på et generelt nivå og burde være av interesse for alle som ønsker å bruke datateknikk av noe omfang.

### Gammelnorskavdelingens arkiv; et seddelarkiv over et avgrenset korpus

Gammelnorskavdelingens arkiv består av 700 000 ordsedler basert mest på litterære tekster, men også noe på diplomer. Sedlene har allminnelig ekserptseddelutforming. Et eksempel er vist i fig. 1.

1305-09-25 S,C  
Nidaros

*sunr, m  
Biarne Erlings son ns*

(1) Biarne Erlings son j. Biarkrœy. Erlingr amunda son. Snare  
(2) aslaks son. ok hallsteinn Thorleifs son. Sænnda allum monnum  
j. ve(3)radale þeim sem / þetta bref sea æða hœyra. Quediu Guds  
ok sína. Af þui at (4) sua profædezst firir oss at sinni æftir  
þvi sem þeir baro Thorsteinn (5) bonde armannz son / ok Haralldr.  
at þæssir mænn er her næmfnazst j. [gengu í vorzlu fyrir Áslák].

DN III 56 Nidaros 1305.

Fig. 1 Original seddeltekst fra gammelnorskarkivet

Seddel fra gammelnorskarkivet

Nummerfelt:	<input type="text" value="1303 09 25"/>	<input type="checkbox"/>	Oppslagsord	<input type="text" value="sunr"/>
Språkform:	<input type="text"/>		Gram. oppl:	<input type="text" value="m"/>
Utferdersted:	<input type="text" value="Nidaros"/>		Ordform:	<input type="text" value="son"/>
Miljø:	<input type="text" value="S,C"/>		Gram. oppl:	<input type="text" value="ns"/>
Første linje:	<input type="text" value="1"/>			
Seddeltekst:	<div style="border: 1px solid black; padding: 5px;"> <p>(1) Biarne Erlings son j. Biarkrœy. Erlingr amunda son. Snare (2) aslaks son. ok hallsteinn Thorleifs son Sænnda allum monnum j. ve(3)radale þeim sem / þetta bref sea æða hœyra. Quediu Guds ok sína. Af þui at (4) sua profædezst firir oss at sinni æftir þvi sem þeir baro Thorsteinn (5) bonde armannz son / ok Haralldr. at þæssir mænn er her næmfnazst j. [gengu í vorzlu fyrir Áslák].</p> </div>			
Kilde:	<input type="text" value="DN"/>	<input type="text" value="3"/>	<input type="text" value="56"/>	
	(verk)	(bind)	(tekstnr)	

Skjelett-korr. OK

Fig 2 Seddelen skrevet inn i registreringsskjemaet

Før man starter registreringen av en slik seddelmasse, er det viktig å foreta en grundig analyse av sedlenes form og innhold. Det er vel og bra å finne hvilke felter som er nødvendige. Men det er like viktig å undersøke hvilke data som deles, dvs. er de samme på mange sedler. For å få dette klart frem lønner det seg å bruke et datamodelleringsverktøy. Vi har valgt NIAM (Nijssen 1988) og vil forlange at alle databasene i prosjektet beskrives i dette systemet. Se



også (Hjort 87) for bruk av formelle datamodelleringsverktøy innen leksikografi. I det følgende blir det gitt en noe uformell analyse av gammelnorsksedlene.

En kontekst er brukt til 30-40 sedler. Det er derfor bare rundt 30 000 forskjellige kontekster blant de 700 000 sedlene. Det er altså et stort rasjonaliseringspotensiale. Videre finner man ut at en slik seddelkontekst er entydig bestemt ut fra den trykte kildens tittel, hvilket bind og hvilken side konteksten er hentet fra, samt hvilken linje på denne siden konteksten begynner med. Tallene i parenteser i konteksten angir linjenumrene i den trykte kilden (se fig. 1). Ved å legge inn en test i databaseprogrammet er det dermed mulig å sikre seg mot at en kontekst blir skrevet mer enn en gang. Innskriverne vil først gå gjennom sedlene og bare skrive av kontekstene. Når dette er gjort, kan den sammenhengende teksten i et verk rekonstrueres ved at kontekstene sorteres etter sidenummer og linjenummer. I databasen vil imidlertid kontekstene få en fortløpende intern nummerering i henhold til den rekkefølgen de registreres i. I tillegg denne tabellen lages det en tabell over kildehenvisninger (som kunne rasjonaliseres ytterligere) som vist i tabell 4 under.

Tabell 1:

Gram.oppl	Ordkl.	Ordform	Norm.form
.	.	.	.
ns	m	son	nei
ns	m	sunr	ja
.	.	.	.

Tabell 2:

Norm.form	Ordform
sunr	son
sunr	son
.	.
.	.

Tabell 3:

Ordform	Kontekstnr	B.pos.	S.pos.
son	12342	122	124
son	845	17	19
.	.	.	.
.	.	.	.

Tabell 4:

Kontekstnr	Verk	Tekst	F.linje
1	Thomas	1	1
12342	DN 3	56	8
.	.	.	.
.	.	.	.

Fig. 3 Skissemessig fremstilling av noen av tabellene i gammelnorskdatabasen

Kontekstene på sedlene er hentet fra trykte utgaver, så man kan spørre om hvorfor vi ikke rett og slett skriver av de trykte kildene. Svaret er at det er lagt et stort arbeid i å justere kontekstene i henhold til originalmanuskriptene.

Når alle seddelkontekstene til et verk, f.eks. en saga, er skrevet inn, starter innskrivningen av ordopplysningene på de enkelte sedlene. Disse står i øvre høyre hjørne (se fig.1 og 2) og er "koblet" til konteksten ved at ordformen er understreket (med rødt på originalsedlene). Denne koblingen ønsker vi å bevare. I innskrivings skjemaet (fig. 2) markerer innskriverne ordformen i teksten på skjermen på Macintosh-manér. I databasen er det så en tabell som til hver ordform forteller hvor ordformen forekommer (tabell 3 i fig. 3). Her lagres kontekstnummeret og start- og sluttposisjon i konteksten. Ved å koble tabell 3 og tabell 4 i fig. 3 kan man få frem verk, bind, side og linje som i en alminnelig konkordans, men også

ordformens nøyaktige posisjon i teksten. De grammatiske opplysningene og forbindelsen ordform og normalisert form kan lagres som vist i tabell 1 og 2. i fig. 3.

Registreringen av sedlene gir en database over sedlene, altså en elektronisk variant av det tradisjonelle arkivet. I tillegg åpnes en rekke andre innganger: søking etter ordformer etter grammatiske opplysninger samt fritekstsøking. Det fullstendig nye er at tekstene er koblet til sin egen ordliste. På dertil egnede datamaskiner (Mac, PC med Windows) kan man nå markere et ord og umiddelbart få vist frem seddelinformasjonen om ordforekomsten. Denne såkalte hypertekst-muligheten kommer rett og slett som en bivirkning av registreringen av sedlene.

### Nynorskavdelingens arkiv; et mangeartet seddelarkiv

Nynorskavdelingens arkiv er svært forskjellig fra gammelnorskavdelingens både når det gjelder størrelse (3,2 millioner sedler) og innhold. Sedlene er basert på ekserpter fra litteratur, presse og eldre ordbøker, men også på opplysninger om muntlig og skriftlig bruk av ord fra informanter spredd rundt i Norge. Det er således ikke mulig å velge samme løsning som for gammelnorsk. Det er også liten hensikt i å starte på A og skrive av hele arkivet. I stedet har vi valgt å behandle sedlene i henhold til innhold og opprinnelse. Som nevnt, faller sedlene i tre grupper: sedler med ekserpter fra ordbøker, sedler ekserpter fra litterære kilder samt målføre- og informantsedler. Vi har begynt med sedlene med ekserpter fra ordbøkene til Aasen og Ross. Disse finnes allerede som deler av et grunnmanuskript for Norsk ordbok. Manuskriptet er fra 1930-årene og er på 13 000 maskinskrevne sider. Ordboksteksten blir for tiden tagget på en tilsvarende måte som *Oxford English Dictionary* (OED)(Tompa 1988) og skrevet inn i med et vanlig tekstbehandlingsprogram. Dette er tilstrekkelig siden teksten ikke inneholder posisjonskoblinger og kopier av like tekstfragmenter slik som gammelnorskarkivet (og alle andre seddelsamlinger basert på sammenhengende tekster). Et utsnitt av den taggedede teksten er vist i fig 4.

```
/OPPF hamp-åker /GRMr m, /DEFI åker der det veks hamp, /KjFd Å. /OPPF hamra
/HONr I. /GRMr v. /BØYr (a) /DEFI arbeida på med hamar. /MÅLF òg "hambra"
/HMFm (sumst.), /KJFd Å. /OPPH Jfr. hamar I. Jfr. nisl.
```

Fig 4 En bit av det taggedede manuskriptet til *Norsk ordbok*.

Når ordboksteksten er ferdig innskrevet, blir den analyseres og senere lagt opp som en database. Til innskrivningen er det i skrivende stund engasjert en forskningsassistent og fire personer på sysselsettingsmidler.

Det neste steget i registreringsarbeidet er å gå løs på registreringen av informant-sedlene. Hvordan sedlene med ekserptene fra de litterære kildene skal behandles, er ennå ikke bestemt. Her gjelder det å tenke som om oppgaven skulle gjøres på nytt. Det er mulig at den beste løsningen vil være å lese optisk hele referansebibliotektet og analysere dette maskinelt.

Det endelige målet for både nynorsk- og bokmålsdelen av prosjektet er å lage en autoritativ database over det norske skrift- og talemålet som støtte for ordboksproduksjon, men også som et hjelpemiddel for annen språkforskning.

## Fritekstsøking kontra databaser

For å gi våre avdelinger en viss følelse av hvilke muligheter datateknikken kan gi, har vi tatt for oss de elektronisk tilgjengelige versjoner av manuskriptene til *Bokmålsordboka* og *Nynorskordboka*. De manuskriptene vi har fått, er stort sett løpende tekst full av settekode, men det har vært foretatt noe analyse og feltinndeling. Videre er det i manuskriptene en del tilleggsinformasjon om godkjente sideformer. Disse opplysningene er ikke med i de trykte utgavene. Vi har analysert tekstene og tilrettelagt dem for fritekstsøk. I tillegg har vi laget en sterkt redigert versjon som danner basis for våre databaser.

Det å lage en fritekstsøkeversjon av manuskriptene innebærer en viss tagging. Dette ble i vårt tilfelle gjort automatisk. Resultatet ble et system der man kan søke etter vilkårlige tekststrenger i ordbøkene og få vist resultatet i et "pent" format på skjermen. Vi har her brukt programvare utviklet ved University of Waterloo, Canada, til bruk i forbindelse med OED (Tompa 88). Det var ganske enkelt å få laget søkeversjoner av de norske ordbøkene. Metoden må sies å være lite kostnadskrevende hva angår informatikerinnsatsen. Man får altså et slagkraftig søkeverktøy uten stor innsats. Men resultatet er lite fleksibelt i den forstand at det ikke er mulig å få skrevet ut statistikker og tabeller uten innsats fra profesjonelle programmere.

Å legge ordbøkene opp som databaser, dvs. plassere opplysningene i tabeller som vist for gammelnorsk, er mer krevende. Tekstene er ordboksmanuskripter. De er ment å skulle leses som løpende tekst. De inneholder også en del inkonsistens både på formatplan og i det innholdsmessige. Noe annet er ikke å vente, og som manuskripter betraktet er de svært ryddige. Men de interne særegenhetene gjør det vanskelig å omforme manuskriptene til fullverdige databaser. Vi har likevel kommet langt i databaseoppbyggingen, men ideelt sett burde en leksikograf gå gjennom databasene og rette inkonsistenser. Et annet punkt er hvor langt man kan fjerne seg fra ordboksformen og likevel kunne rekonstruere en ordbok automatisk fra databasen. Fordelen med slike databaser er at det åpnes for å foreta søk og lage tabeller hinsides de mulighetene som ligger i en fritekstdatabase.

Når en ordbok skal legges opp som en tradisjonell database, kan imidlertid selve databaseverktøyet bli en kompliserende faktor. Det synes å være mindre informatikk-krevende å velge en enbrukerløsning på en PC eller Macintosh enn den flerbrukerløsningen vi har valgt. Databasene er lagt opp i INGRES, som er et profesjonelt flerbrukersystem. Fordelen med dette systemet er at mange brukere kan bruke databasene (både lese og skrive) samtidig, samt at ulike brukere kan gis ulike privilegier. Ulempen er at opplegg, drift og vedlikehold av en slik database krever spesialopplært personell. Tross de større informatikkmessige kostnadene mener vi likevel at det er best å velge denne siste løsningen. Det gir oss muligheten til å knytte leksikografene til én database der databasesystemet automatisk kontrollerer tilgangsprivilegier og takler at flere brukere gjør forandringer i databasen samtidig. Orddatabasene kan dermed også bli tilgjengelige for brukere utenfor de leksikografiske avdelingene samtidig som det er mulig å kontrollere adgangen.

## Litteratur

- Hjort, Ebba. 1987. The Danlex-Group, *Descriptive Tools for Electronic Processing of Dictionary Data Lexicographica Series Maior 20* Max Niemeyer Verlag, Tübingen 1987

G.M. Nijssen, T.A.Halpin. 1988. *Conceptual Schema and Relational Database Design*, Prentice Hall, New York 1988

Tompa, F. Wm., D.L.Berg, G.H.Gonnet. 1088 *The New English Dictionary Project at the University of Waterloo*.

Centre for the NOED. University of Waterloo, Waterloo, Canada

Jan Engh

## Leksikografi i IBM Norge

IBM Norge driver den største leksikografiske virksomheten her til lands utenfor universitetene. Arbeidet er samlet om databehandlingsterminologi og allmennspråk - både bokmål og nynorsk. Terminologiarbeidet er viktig for oversettervirksomheten i firmaet. Arbeidet med norsk allmennspråk er en forutsetning for utviklingen av språklige funksjoner i dataprogram. Mens de fleste programverktøyene er utviklet i utlandet, er det materialet de blir brukt på, stort sett lagd i Norge. Derfor vil det meste av framstillingen dreie seg om leksikografisk produksjon. Men først vil jeg nevne kort hva slags hjelpemidler vi har for å *bruke* leksikografisk materiale, hvilke ordbøker og orddatabaser vi har tilgang til, og hva slags verktøy vi har til hjelp i terminologiarbeidet. Deretter kommer jeg til å si litt om de allmennspråklige databasene vi har *lagd*, og å drøfte enkelte språklige problem vi har støtt på under arbeidet. Til slutt vil jeg komme inn på hvordan vi *utnytter* disse databasene.

### 1. Leksikografiske hjelpemidler

Behovet for verktøy til oversetterne<sup>1</sup> har ført til at det er installert maskinleselige ordbøker som en kan slå opp i på skjermen. Når en er inne i en tekst, kan en ved en enkel kommando få opp den ordboksartikkelen en ønsker i et vindu.

Det programmet som er brukt for oppslaget i figur 1, heter WordSmith<sup>2</sup>. Ved hjelp av det kan medarbeiderne slå opp i en rekke maskinleselige ordbøker og ordlister:

- Webster 7. utgave (engelsk ettspråklig ordbok)
- Webster 7. utgave (etymologiske opplysninger)
- *Collins Thesaurus* (engelsk synonymordbok)
- *Roget's Thesaurus 2.* utgave
- IBM (Woods): *Dictionary of Computing*
- IBM (Quade) engelske forkortelser
- Kunnskapsforlaget: *Engelsk-norsk ordbok*
- Kunnskapsforlaget: *Norsk-engelsk ordbok*
- Kunnskapsforlaget: *Fremmedordbok*
- *Bokmålsordboka* (ettspråklig ordbok for norsk bokmål)
- Rommetveit: *På godt norsk* ("fornorskingsordbok" med synonymer)
- IBM Norge: *Synonymordbok* (norsk bokmål)
- TFSNOR (engelsk-norsk dataterminologi)
- forskjellige særtermlister: SAA, OS2, AIX osv.

```

WordSmith Dictionary      ENGNOR: (c) IBM Norge AS      IBM Internal Use Only
      gnostic                      hest
      gns                          invoice
      gnu                          kort
+-----+
!
! go (v)                          !
! gå, dra, dra av sted, reise, ta (et sted
!   hen), begi seg, gå av (om skytevåpen)  !
! lyde, ringe (om klokke), slå (om ur), være i !
!   omløp (om rykte)                    !
! anses for                          !
! ha til formål, nå, føre til, finne sted  !
! lykkes, befinne seg                  !
! gå ut på, foreta, ha til hensikt      !
! ta tilflukt til                      !
! være drektig                        !
!
! go (n)                              !
! gang                                !
! hending                             !
! omstendighet, affære                !
! siste skrik, mote                   !
! energi, mot, pågangsmot, futt       !
+-----+-----More+
go about
go about it
go about your own busi
go a boxing
DICTIONARY: ENGNOR      COMMAND: cms scrncopy
DIMENSION 1: =          DIM 2:          DIM 3: FINDLOG DIM 4:
1=HELP 2=CMNT 3=QUIT 4=FOUND 5,6=SCRDIM 7,8=SCRBOX 9=SAVE 10=CURS 11=FLIP 12=RET

```

Figur 1. Oppslag på skjermen i Kunnskapsforlagets *Engelsk-norsk ordbok*

Disse bøkene er tilgjengelige for alle eller for visse grupper av medarbeidere - alt etter hvilke konfidensielle opplysninger de inneholder, og hvilke rettigheter IBM Norge har til bruken av hver enkelt ordbok (intern bruk i en enkelt gruppe eller avdeling, eventuelt til bruk i hele organisasjonen).

Det er svært ulike ordbøker vi kan gjøre oppslag i på denne måten, og de er av forskjellig opphav. Noen er kjøpt (for eksempel *Bokmålsordboka* og *På godt norsk*). Andre har blitt lagd i huset. Den siste gruppa omfatter både allmennspråk (synonymordbok for bokmål) og terminologi (forskjellige særtermelister og TFSNOR på rundt 15 000 oppslagsord).

Særskilt verktøy kan brukes for å skille ut og arbeide med terminologi: TIC (Terminology Identification using CRITIQUE<sup>3</sup>) går igjennom de engelske utgangstekstene, og stiller opp lister over mulig terminologi som mangler norsk oversettelse, både enkeltord og uttrykk med flere ord. TFS (Terminology Filing System) er en "arbeidsbenk" for terminologen. Begge disse programmene er samlet i ett system for styring av termarbeidet, TMS (Terminology Management System). Resultatet lagres i TFSNOR<sup>4</sup>. I tillegg har IBM Norge mer detaljerte databaser som en kan gjøre intelligente søk i. Ett slikt databaseformat

er LDB (Linguistic Database). Blant annet *Bokmålsordboka* finnes i LDB-format<sup>5</sup>. Dette er viktig i samband med utviklingen av funksjoner som virker på norsk.

## 2. Leksikografisk produksjon

I datalingvistisk forskning og utvikling går IBM gjerne ut i full bredde. Det nytter lite å lage system som når de store analytiske dybder, dersom de bare har et ordforråd på ett hundre ord eller to - både språkvitenskapelig sett og av kommersielle grunner. Alle IBMs funksjoner som virker på naturlige språk, har en høy dekningsgrad.

En nødvendig forutsetning og et grunnlag for utvikling av slike funksjoner er et omfattende leksikon og en fullstendig morfologi. Da IBM Norge begynte å arbeide med datalingvistikk i 1984, fantes det ikke noen maskinleselige ressurser av dette slaget for norsk som var offentlig tilgjengelige. Alt måtte lages fra grunnen av. Resultatet av arbeidet som er lagt ned siden da, er en meget omfattende dokumentasjon av normeringen av moderne norsk allmennspråk, sannsynligvis den mest omfattende som finnes.

### 2.1 Leksikon og morfologi

Grunnstammen i ordsamlingen er to leksikon, ett for bokmål og ett for nynorsk. Ingen av dem inneholder sideformer<sup>6</sup>. Til hvert leksikon er det knyttet en fullstendig og komplett morfologi; fullstendig i den forstand at hvert lemma i leksikonet er tilordnet ett bøyingsmønster i morfologien, komplett siden den dekker alle bøyingsformene til hvert lemma - også dem som vel bare er mulige reint teoretisk (se figur 2).

rik	finne
rike	finner
rikt	finnes
rikere	fant
rikest	funnet
rikeste	funnet
riks	funnets
rikes	funnets
rikts	funne
rikeres	funnes
rikestes	finnende
	finnendes
	finn

Figur 2. Alle former av adjektivet **rik** og verbet **finne**.

Vi kan når som helst kjøre ut fullstendige lister med alle former av alle ord i hvert leksikon, som i figur 3, eller for eksempel ta ut lister over alle ord som bøyes etter samme mønster som i figur 4.

Bokmålmorfologien omfatter 629 bøyingsmønster, nynorskmorfologien 576. I dag<sup>7</sup> omfatter bokmålsleksikonet om lag 79 000 lemma, nynorskleksikonet nærmere 111 000. Selve tallet på bøyingsmønster er litt høyere enn det som er lingvistisk motivert. Det kommer av at formatet i det systemet som skal ta vare på leksikon og morfologi, stiller visse særskilte tekniske krav. Likevel gir tallene en pekepinn om kompleksiteten i moderne norsk morfologi.

gjøremål	800 0	gjøremål
gjøremålet	800 0	gjøremål
gjøremål	800 0	gjøremål
gjøremåla	800 0	gjøremål
gjøremålene	800 0	gjøremål
gjøremåls	800 0	gjøremål
gjøremålets	800 0	gjøremål
gjøremåls	800 0	gjøremål
gjøremålas	800 0	gjøremål
gjøremålernes	800 0	gjøremål

Figur 3. Utsnitt av fil som viser alle former av alle ord i et leksikon. Ordformer i første kolonne, henvisning til bøyingsmønster i andre og oppslagsform i kolonnen lengst til høyre.

```
====>
!...+...1...+...2...+...3...+...4...+...>
```

* * * Top of File * * *		
beltespenne	790	00000
binyre	790	00001
åte	790	00002
nyre	790	00003
sølvspenne	790	00004
skyvelære	790	00005
spenne	790	00006
* * * End of File * * *		
		00007
		00008

Figur 4. Skjerm bilde som viser alle ord i samlingen som bøyes som nyre.

Til grunn for denne ordsamlingen ligger det et omfattende forsknings- og utredningsarbeid som også er utført av og hos IBM Norge, delvis i samarbeid med Norsk språkråd. Dette har vært et nybrottsarbeid - ikke bare fra et kvantitativt, men også fra et kvalitativt synspunkt. Ikke nok med at en omfattende maskinleselig oversikt over norske ord ikke fantes, trykte ordbøker viste seg også å være både upålitelige og ufullstendige. Vi kan snakke om både hvite flekker og gråsoner i trykte ordbøker for norsk<sup>8</sup>. La oss se nærmere på situasjonen for bokmålet. Først forhold som gjelder normering: Ikke noen trykt ordbok gjengav gjeldende rettskrivningsvedtak. Det har vært nødvendig å sjekke alle data mot årsmeldinger og liknende fra Norsk språkråd (i den grad det gjaldt gamle vedtak), over telefonen med medarbeiderne i Norsk språkråd (ved nyere vedtak). Men et hovedproblem er selvsagt å vite hva som er galt, og som bør kontrolleres i ordbøkene.

Det kan by på problem når en skal tolke gjeldende normal. Mens arbeidet ennå omfattet både hoved- og sideformer, var ett slikt problem hva en kunne slutte om tillatte sideformer til sammensatte ord der ett av leddene hadde en sideform. For eksempel **vinne** og **overvinne**:



<b>vinne</b>	vant/[vann]
<b>overvinne</b>	overvant/[?overvann]

Det finnes mange unormerte ord, særlig fremmedord, der vi har vage intuisjoner om bøyningsformer. Hva heter for eksempel flertall bestemt form av **mile** og **royalty**? Etter betenkningstid har Norsk språkråd kommet til at det skal hete:

<b>mile</b>	milen	miles	milene
<b>royalty</b>	royaltyen	royalties/royaltyer	royaltyene

Vi har også oppdaget en rekke forhold som ingen hadde tenkt alvorlig over, før datamaskinen krevde et nøyaktig svar. Hva heter for eksempel perfektum partisipp av **briste**? Eller presens partisipp av **be** og **bede**. Og er "beding" en form av **be**<sup>9</sup>? Slik de er nå, må vi regne med at samlingene er så godt som helt *riktige*.

La oss så se på det vi kan kalle deskriptive forhold; i praksis vil det gjerne si om ord har komplette bøyingsmønstre. Hver ordklasse har visse kategorier som det er vanskelig å si er komplette eller ikke ut fra semantiske forhold. Et godt eksempel er attributive former av verb. Mengden av transitive verb og verb med attributive former er ikke identiske, og i praksis er en nødt til å ta stilling til hvert enkelt verb for seg. Ofte kommer en også opp i normative problem, siden det kan være tvil om hvilket uttrykk en attributiv form skal ha. Hva heter for eksempel de attributivt bøyde formene av perfektum partisipp av **briste** og **sprekke**?

Flertall av substantiv kan være vanskelig. Vi kan stille opp en skala av substantiv, der eksemplene i ytterendene for de fleste språkbrukere enten har eller ikke har flertallsformer<sup>10</sup>.

<b>fot</b>	føtter
<b>seneskjedefetnesle</b>	(?)seneskjedefetnesler
<b>godhet</b>	?godheter
<b>lykke</b>	?lykker
<b>rød</b>	*røder
<b>godfot</b>	*godføtter

Et tredje eksempel er gradbøyning av adjektiv. Mellom

<b>snill</b>	snillere	snillest
og		
<b>faglig</b>	*fagligere	*fagligst

finnes det en rekke adjektiv der det er mer eller mindre vanskelig å godta egentlig gradbøyning<sup>11</sup>. For eksempel

<b>kompleks</b>	(?)kompleksere	?kompleksest
<b>smellfeit</b>	?smellfeitere	?smellfeitest

Men kompletthet er ikke bare et semantisk problem. Et eksempel på det er genitivsformene av adjektiv. Det er tvilsomt om visse adjektiv kan forekomme i genitiv entall intetkjønn positiv, og de kan ikke opptre i genitiv superlativ sterk form entall.

rik riks rikes ?rikts rikeres \*rikests rikestes

Under alle omstendigheter er det helt klart at dette er helt marginale ordformer<sup>12</sup>, men grunnen til defekten må ha med fonotaktiske/grafotaktiske forhold å gjøre<sup>13</sup>.

Det er alltid en vurderingssak om en skal ta med slike former. Vanligvis lar en være å ta stilling til spørsmål av dette slaget overhodet, men når en har å gjøre med den pedanten som en datamaskin er, må en ta stilling. Og det har vi gjort. Også på dette området er sannsynligvis samlingene våre de mest pålitelige og omfattende som finnes i dag. Men det betyr naturligvis ikke at de ikke kan bli bedre.

## 2.2 Annen produksjon

Den mest konvensjonelle delen av den leksikografiske virksomheten er redigeringen av synonymordbøker. IBM Norge har lagd en egen synonymordbok for bokmål på rundt 17 500 oppslagsord og en for nynorsk på om lag 27 000 oppslagsord er snart ferdig. Den nynorske synonymordboka har blitt "optimalisert", det vil si at alle ord og uttrykk i definisjonsdelen av artiklene har blitt vurdert som mulige oppslagsord. Alle oppslagsord og ord i definisjonene dekkes av leksikon og morfologi i hver sin målform.

Et eget prosjekt består i å merke alle ord i leksikon etter de syntaktiske og stilistiske egenskapene deres. Dette er opplysninger som kommer til nytte i arbeidet med å lage en syntaks for norsk<sup>14</sup>. Figur 5 viser et lite utsnitt av lista over merker som vi til nå har hengt på ord alt etter de syntaktiske eller stilistiske egenskapene deres.

```
NOB      ODTABLE A1 V 80 Trunc=80 Size=431 Line=345 Col=1 Alt=0
====>
!...+...1...+...2...+...3...+...4...+...5...+...6...+...7..

420 0 ENEADVPS adverb eller adjektiv som bare kan stå som ADVP aleine      00343
      i en setning                                                            00344
427 0 MAAL ord som uttrykker mål, f eks KILOMETER, KM, GRAM                 00346
428 0 MENGDE uttrykker mengde                                               00348
429 0 FARGE ord som viser til en farge, f eks GUL, HVITT, KARMOSIN          00350
430 0 TVGPRP preposisjon i "transitiv verbalgruppe",                       00352
*      "non flippable" dvs kan IKKE sprette over til høyre for             00353
*      verbet, "på" i "tenke på"                                           00354
*      "Han tenkte på saka."                                               00355
*      "**Han tenkte saka på."                                             00356
431 0 TVGPRPF preposisjon i "transitiv verbalgruppe"                       00358
*      "flippable" dvs KAN sprette over til høyre fra verbet,             00359
*      "på" i "ta på"                                                      00360
*      "Han tok på frakken"                                                00361
*      "Han tok frakken på"                                               00362
432 0 VPART adverb som verbalpartikkel i løst sammensatt verb,             00364
*      "non flippable" dvs kan ikke sprette over til høyre fra           00365
*      verbet, "an" i "legge an", "over" i "ta over"                     00366
```

Figur 5. Utsnitt av liste over merker som viser til syntaktiske eller stilistiske trekk ved ord.

### 2.3. Verktøy

Til hjelp i arbeidet med leksikon, morfologi og synonymordbøker har vi verktøy som kontrollerer at den syntaksen som systemet krever<sup>15</sup>, blir fulgt (jf figur 6).

```

DIAG2   MESSAGES A1 V 80 Trunc=80 Size=5 Line=0 Col=1 Alt=0
====>
!...+...1...+...2...+...3...+...4...+...5...+...6...+...7..

* * * Top of File * * *
1* ERROR-- NO MATCH FOR FORM ONE.          00000
  CLASS: 335                               00001
  BASE WORD: ettergi                       00002
  FORM ONE: e+%e                           00003
  STEM / FAULTY MATCH: ett ergi            00004
* * * End of File * * *                    00005

```

Figur 6. Feilmelding som viser at verbet **ettergi** i leksikonet ikke kan bøyes etter mønsteret 335. Denne regelen (bøyningsmønsteret) stiller formelle krav (hvilke egenskaper som de siste fire bokstavene i verbet skal ha) som sekvensen "ergi" ikke oppfyller.

For å gjøre enkeltoppslag og se hvilke grammatiske trekk som særmerker en gitt ordform i en ferdigbygd leksikonmodul, kan vi bruke editoren ODL ("Optimised Dictionary") for oppslag i leksikon). Et eksempel er vist i figur 7.

```

hvalfangstens                               OD/morphology codes
====>
!...+...1...+...2...+...3...+...4...+...5...+...6...+...7..

* * * Top of File * * *                    00000
hvalfangst(NOUN HAN ENT BST GEN)          00001
* * * End of File * * *                    00002

```

Figur 7. Editoren ODL. Morfologiske trekk ved "hvalfangstens"

ODL har også en funksjon som gjør at en kan rette, ta bort eller legge til nye trekk. For å gjøre oppslag og se hvilke spesielle syntaktiske trekk som særmerker ett gitt lemma i en ferdigbygd leksikonmodul, kan vi bruke editoren OD (Optimised Dictionary), jf figur 8.

```

ta                                     OD/grammar codes
====>
!...+...1...+...2...+...3...+...4...+...5...+...6...+...7..

* * * Top of File * * *                00000
ta(VERB V2 (TVGPRP over) (TVGPRPF på) (VPART (fram opp))) 00001
* * * End of File * * *                00002

```

Figur 8. Editoren OD. Syntaktiske trekk ved TA (ufullstendig).

På samme måte som ODL har OD også en addendafunksjon, som gjør at en kan rette, ta vekk eller legge til nye trekk, for eksempel mens en er i ferd med å skrive regler for å beskrive et syntaktisk delsystem. Det er naturligvis også mulig å ta ut lister av ord som er kjennetegnet ved samme syntaktiske eller stilistiske trekk (jf figur 9 og 10).

```

                                ODLISTQ version 1.00

For language NOB

File to use as input
====> NOB ODDATA *

Part of speech(s)
====> ADJ

Feature(s)
====> bareattr

Matching by (choose only 1):
====>          all part of speech match and all features match
====>          all part of speech match and any features match
====>          any part of speech match and all features match
====> x        any part of speech match and any features match

ENTER to begin processing   PF1 Help   PF2 Quit   PF3 Quit

CMS command

```

Figur 9. Skjerm bilde som en fyller ut for å lage en liste over alle ord av en bestemt ordklasse med samme syntaktiske eller stilistiske egenskap. Her: Adjektiv som er kjennetegnet ved merket BAREATTR (det vil si at de ikke kan stå i predikativ stilling i en setning).

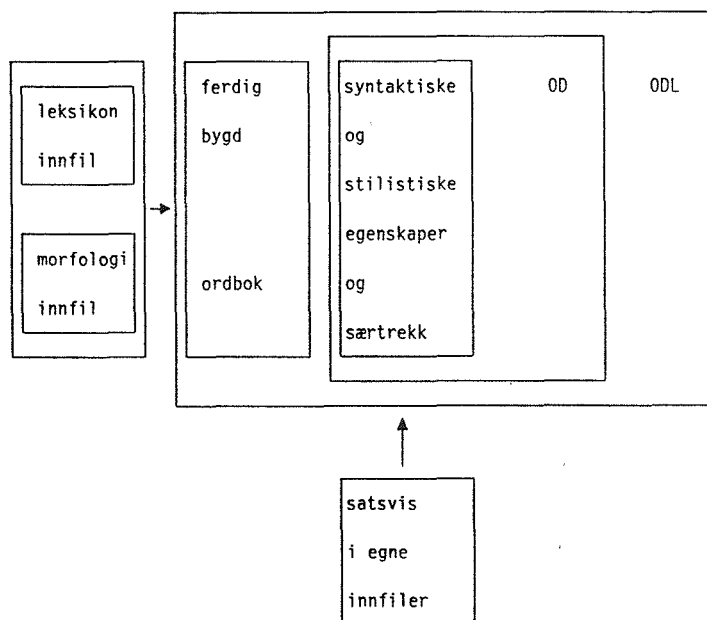
```

NOB      LISTOUT  A1  V 80  Trunc=80 Size=64 Line=3 Col=1 Alt=18
====>
!...+...1...+...2...+...3...+...4...+...5...+...6...+...7..
begge:                                       00009
  begge(ADJ BAREATTR)                       00010
beviselig:                                  00011
  beviselig(ADJ KANADV ENEADVPS BAREATTR)   00012
bevislig:                                   00013
  bevislig(ADJ KANADV ENEADVPS BAREATTR)   00014
bidig:                                      00015
  bidig(ADJ BAREATTR)                       00016
diverse:                                    00017
  diverse(ADJ BAREATTR)                     00018
formentlig:                                00019
  formentlig(ADJ KANADV ENEADVPS BAREATTR) 00020
godeste:                                    00021
  godeste(ADJ BAREATTR)                     00022
havsens:                                    00025
  havsens(ADJ BAREATTR)                     00026

```

Figur 10. Utsnitt av lista over alle adjektiv med merket BAREATTR.

Dessuten går det an å endre og legge til data satsvis, jamfør skissen i figur 11.



Figur 11. Hvordan en i prinsippet kan legge inn data og kikke på hvilke morfologiske og syntaktiske eller stilistiske trekk som kjennetegner ord(former) i ordsamlingen.

### 3. Sluttproduktet

Hva brukes så de allmennspråklige ordsamlingene til? Det finnes en rekke bruksområder: Vi kan tenke oss hele program som virker på norsk språk, for eksempel program for automatisk oversettelse til eller fra en norsk målform. Men den viktigste anvendelsen er som grunnlag for *funksjoner* som virker på norsk språk, altså som deler av andre brukerprogram. Og selv om det finnes språklige funksjoner i for eksempel søkesystem, så er anvendelsesområdet for ordsamlingene først og fremst IBMs egne tekstbehandlingsprogram<sup>16</sup>. Leksikon og morfologi danner blant annet basis for funksjonen *bøyningsopplysninger* og for grunnordlister som igjen styrer stavekontroll og automatisk orddeling av kjente ord.

```

VISFRAM RFT      A1                      STAVING                      Side 1
** Ordet er feilstavet
==>
<---+---1---+---2---+---3---+---V---+---5---+---6---+---7---+--->
Har du en leiekontrakt som ikke er tidsbegrenset (vanlig leiekontrakt),
og du blir sagt opp, kan du få oppsigelsen vurdert av domstolene.
Hovedregelen er at domstolene kan sette oppsigelsen til side hvis den
mangler saklig grunn, eller vil virke urimelig overfor leietakeren.
Det er den begrunnelsen utleieren gir for oppsigelsen, som blir vurdert
av rette +-----+
      Hvi ! begrunnelsen ! ppsigelsen er urimelig overfor leietakeren,
kan rett ! begrunnelse ! elsen er ugyldig , selv om det er saklig
grunn fo ! begrunnelsene ! opp fra leieavtalen og mener at oppsigelsen
ikke fyl ! grunnens ! gyldighet, må du sende stevning til retten
innen 30 ! grunnens ! ppsigelsen.
      +-----+

----- Dokumentslutt -----

PF 1=Hjelp      2=Innskyt      3=Avslutt      4=Koder      5=GSøk      6=Sjekk
PF 7=Tilbake    8=Frem          9=Blokk       10=Bøyning  11=Neste    12=Synonymer

```

Figur 12. Eksempel på bruk av stavehjelp i DW370. Forslag til hva **begrunnelsen** skal være.

```

VISFRAM RFT      A1                      STAVING                      Side 1
** Velg hvilket ord som skal utvides og trykk på ENTER
==>
<---+---1---+---2---+---3---+---V---+---5---+---6---+---7---+--->
Har du en leiekontrakt som ikke er tidsbegrenset (vanlig leiekontrakt),
og du blir sagt opp, kan du få oppsigelsen vurdert av domstolene.
Hovedregelen er at domsto +-----+ sigelsen til side hvis den
mangler saklig grunn, ell ! Verb: få ! lig overfor leietakeren.
Det er den begrunnelsen u ! Adjektiv: få ! ppsigelsen, som blir vurdert
av retten. +-----+
      Hvis retten finner at oppsigelsen er urimelig overfor leietakeren,
kan retten avgjøre at oppsigelsen er ugyldig , selv om det er saklig
grunn for den. Blir du sagt opp fra leieavtalen og mener at oppsigelsen
ikke fyller lovens krav til gyldighet, må du sende stevning til retten
innen 30 dager fra du fikk oppsigelsen.

----- Dokumentslutt -----

PF 1=Hjelp      2=Innskyt      3=Avslutt      4=Koder      5=GSøk      6=Sjekk
PF 7=Tilbake    8=Frem          9=Blokk       10=Bøyning  11=Neste    12=Synonymer

```

Figur 13. Eksempel på bruk av funksjonen bøyning i DW370 (1). Ordformen **få** faller inn under to oppslagsord som hører til to ulike ordklasser.

VISFRAM RFT A1 STAVING Side 1  
Linje 0

```

===>
<---+---1---+---2---+---3---+---V---+---5---+---6---+---7---+--->
Har du en leiekontrakt som ikke er tidsbegrenset (vanlig leiekontrakt),
og du blir sagt opp, kan du få oppsigelsen vurdert av domstolene.
Hovedregelen er at domsto +-----+
mangler saklig grunn, ell ! Bøying av ordet "få":          !
Det er den begrunneisen u ! Verb:                           !
av retten.                ! aktiv                          !
    Hvis retten finner a !      infinitiv                  få       !
kan retten avgjøre at opp ! presens                    får       !
grunn for den. Bli du sa ! preteritum                  fikk      !
ikke fyller lovens krav t ! perfektum partisipp   fått      !
innen 30 dager fra du fik !      i perfektum      fått      !
                                samsvar                  !
                                !      intetkjønn            fått      !
-----+-----+----- !      han- el. hunkjønn   fått      !
                                !      flertall el. svak bøyn. fåtte    !
                                !      presens partisipp   fående    !
                                ! S-former                  !
                                !      infinitiv          fås, fåes !
                                !      presens            fås, fåes !
                                !      presens partisipp   fåendes   !
                                !      imperativ          få         !
-----+-----+-----+
PF 1=Hjelp      2=Innskyt  3=Avslutt  4=Koder    5=GSøk    6=Sjekk
PF 7=Tilbake    8=Frem      9=Blokk  10=Bøying  11=Neste  12=Synonymer

```

Figur 14. Eksempel på bruk av funksjonen bøying i DW370 (2). Bøyningsmønsteret for verbet FÅ som ordformen få tilhører.



DEMON	RFT	A1	STAVING	Side 1
				Linje 2
====>				
<---+---1---+---2---+---3---+---V---+---5---+---6---+---7---+--->				
og du blir sagt opp, kan du få oppsigelsen vurdert av domstolene.				
Hovedregelen er at domstolene kan sette oppsigelsen til side hvis den mangler				
Det er !	!	Synonymer for ordet "sette":	!	!
av rett !	!	Verb:	!	!
Hv !	!	- stille, plassere	!	!
kan ret !	!	- plante	!	!
grunn f !	!	- satse, vedde	!	!
ikke fy !	!	- bringe, føre	!	!
innen 3 !	!	- lage, danne	!	!
!	!	- avtale, beramme	!	!
Vi gjer !	!	<seg>	!	!
derfor !	!	- ta plass	!	!
at du s !	!	- si sammen	!	!
halde t !	!	- feste seg	!	!
set fra !	!	<etter>	!	!
busest !	!	- forfølge	!	!
gjerast !	!	<i gang>	!	!
!	!	- begynne, starte, gå i gang, sette i, klemme på, kjøre i vei	!	!
Este ar !	!	- iverksette, sette i verk, sette ut i livet	!	!
entre e !	!	- få til å starte	!	!
de Madr !	!	<til>	!	!
un dicc !	!	- omkomme	!	!
informa !	!	<inn>	!	!
verific !	!	- gni inn, stryke på	!	!
como he !	!	- si på, senke seg, falle på	!	!
PF 1=Hj !	!	<seg inn i>	!	!
PF 7=Ti +	!		!	!

Figur 15. Eksempel på oppslag i synonymordboka (bokmål) i DW370.

Forskjellige bruksområder stiller ulike krav til det språklige grunnlaget, og dette bringer oss tilbake til leksikografisk produksjon. Ideelt sett er en nødt til å lage egne ordlister til ulike behov. I stedet for å bruke en og samme database må en gjøre utvalg til særskilt bruk. Til grunn for en stavekontroll trenger en for eksempel en liste av bare riktige ordformer. I et system med en språklig komponent for søking i tekst, derimot, trenger en også å ta med visse uriktige ordformer i lista. Det kan dreie seg om tidligere eller regionale varianter. Dette er spesielt viktig når man har med et språk som norsk å gjøre, der rettskrivningsendringene har vært hyppigere og mer radikale enn i de fleste språksamfunner. Et eksempel som viser de forskjellige kravene som ulike brukerprogrammer krever, er ETTER, som er riktig, og som må være den eneste formen som en stavekontroll bygger på, og EFTER, som også må være med i forbindelse med et søkesystem. Hvis ikke, får en ikke med både "etter" og "efter" i eldre tekster eller tekster som ikke følger bokmålsnormen. Søkert blir ikke fullstendig, og kvaliteten på programmet dårlig. Akkurat som når en bygger opp allmennspråklige databaser, er tradisjonelt leksikografisk arbeid en forutsetning for utvikling av språklige funksjoner i brukerprogrammer.

## Noter

1. Opptil 40 oversettere kan være i arbeid på samme tid. I IBM Norge blir det oversatt store mengder av tekniske tekster fra engelsk til norsk: I 1990 hele 4 501 sider med trykt dokumentasjon og 256 362 linjer med maskinkode, "machine-readable information". Altså noe rundt 9 000 sider i alt.
2. Det er et internt systemprodukt som virker på stormaskiner under operativsystemet VM.
3. CRITIQUE er et program for grammatisk og stilistisk analyse av engelsk.
4. Terminologen har også mulighet for å slå opp i IBMs termdatabaser for andre språk, f eks dansk og svensk.
5. Se Baustad 1992.
6. Tidligere opererte vi med sideformer for bokmål, men det har vi sluttet med av praktiske grunner (for å kunne behandle flere lemma som virkelig er forskjellige i stedet), dels av kommersielle grunner: Ordsamlingene er jo utgangspunkt for språklige funksjoner i programvare, og kundene ønsket funksjoner som bygde på læreboknormalen, altså uten sideformer.
7. 30.05.1991.
8. Jf Akø 1992.
9. Dette har vi ryddet opp i med velvillig hjelp fra Norsk språkråd.
10. Og tilsvarende for manglende entallsformer av mange substantiv, naturligvis.
11. Her ser jeg bort fra gradbøyning med "mer" og "mest", som er uten interesse i denne sammenhengen.
12. Jf figur 2.
13. Jf adjektiv som *nøye*, der en form genitiv entall intetkjønn positiv, "nøyes", i og for seg er fullt akseptabel.
14. Syntaksen er i sin tur viktig for eksempel for utviklingen av system for automatisk oversettelse med en norsk målform som utgangsspråk og/eller målspråk.
15. Altså ikke syntaks i språkvitenskapelig forstand.
16. "Skriveassistent", DisplayWrite 5 (DW5) for PS-verdenen, DisplayWrite 400 (DW400) for de mellomstore maskinene og DisplayWrite 370 (DW370) for stormaskinene.

## Litteratur

- Akø, J.O. 1992. Gråsoner i norske ordbøker. I: Fjeld, R.V. (red.). 1992. *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden, 28.-31. mai 1991*. Oslo.
- Baustad, J. 1992. Automatisk analyse av maskinleselige ordbøker til bruk i en orddatabase. I: Fjeld, R.V. (red.). 1992. *Nordiske studier i leksikografi. Rapport fra Nordisk konferanse i leksikografi i Norden, 28.-31. mai 1991*. Oslo.

Jostein Baustad

## Automatisk analyse av maskinleselige ordbøker til bruk i en orddatabase

Denne artikkelen skisserer hvordan automatisk analyse av maskinleselige ordbøker til bruk i en orddatabase kan foregå. To typer orddatabaser med ulik detaljeringsgrad beskrives, og to ganske forskjellige framgangsmåter skisseres. Den ene typen orddatabase egner seg godt for vanlige ordboksoppslag, det vil si ved at brukeren oppgir et oppslagsord og får presentert informasjon om oppslagsordet. Den andre typen orddatabase egner seg for avanserte databasesøk blant ordene i databasen. Et eksempel på et slikt søk kan være "finn alle substantiv som stammer fra italiensk og har med domenet musikk å gjøre".

### 1 Mål og hensikt

Denne artikkelen skisserer en framgangsmåte som er benyttet for å legge *Bokmålsordboka* fra Universitetsforlaget inn i en database. Tilsvarende framgangsmåte eller deler av den er også benyttet for andre ordbøker, blant disse *Norsk-Engelsk ordbok* og *Fremmedordboka* fra Kunnskapsforlaget.

Ei ordbok i en database kan brukes på mange måter. To av disse er interessante i denne sammenheng :

- \* Ordboka kan brukes som ei tradisjonell ordbok, det vil si ved at brukeren oppgir et oppslagsord og får presentert all informasjon om dette ordet som ligger i basen. Dersom brukeren er inne i for eksempel et tekstbehandlingsprogram, kan han ved hjelp av et tastetrykk slå opp det ordet som markøren befinner seg i. Vi bruker et program kalt *WordSmith* til denne typen oppslag.
- \* Ordboka kan brukes til å utføre avanserte databasesøk, det vil si finne alle oppslagsord i ordboka som tilfredsstiller de kriterier brukeren oppgir. Et eksempel kan være å finne alle intetkjønnsord som stammer fra italiensk, et annet eksempel kan være å finne alle ord som har med musikk å gjøre. Til denne typen søk bruker vi et program kalt *Lexical Query Language* (LQL).

For å kunne bruke et av programmene ovenfor (eller andre programmer eller verktøy for den saks skyld) er det en forutsetning at ordboka er i et format som programmet kjenner. Det vil si at *WordSmith* og *LQL* trenger ulik kunnskap om ei ordbok. Mens *WordSmith* bare trenger å vite oppslagsordet samt hva som skal presenteres når dette ordet slås opp, så trenger *LQL* detaljert kunnskap om ordklasse, ordhistorie, domener, definisjoner og så videre. *WordSmith* benytter derfor et format kalt *Dictionary Access Method* (DAM), mens *LQL* arbeider med ordbøker i *Lexical Data Base-format* (LDB). Disse formatene beskrives nærmere seinere.

Når ei ordbok skal legges inn i en database, er det hovedsaklig to måter det kan gjøres på. Den ene måten er å bygge opp ordboka manuelt fra grunnen av. Dette er selvsagt veldig tidkrevende, men en kan selv bestemme hvilke opplysninger som en ønsker å ha med i ordboka. Den andre måten baserer seg på mest mulig automatikk ved at en lager et program som leser og analyserer datafiler fra ordboka, og så legger inn hvert oppslagsord med tilhørende opplysninger i databasen. Denne måten krever at man har tilgang til datafiler fra ordboka. Slike filer kan vanligvis skaffes fra forlaget som har utgitt ordboka ved at man kjøper rettigheter til å bruke den i maskinleselig utgave. Det er denne siste måten som beskrives i denne artikkelen.

Datafilene fra forlaget, heretter kalt kildefiler, er vanligvis ikke i det formatet man ønsker. Derfor må kildefilene konverteres til et format som gjør at ordboka kan brukes slik en vil. Det er naturligvis meget viktig å vite hva en ønsker å bruke ordboka til før en utfører konverteringen. Hvis ordboka kun skal brukes som ei tradisjonell ordbok, er det unødvendig å utføre den krevende konverteringen til LDB-format, men dersom en ønsker å gjøre avanserte databasesøk i ordboka, er ikke DAM-formatet tilstrekkelig.

## 2 Ordboksformatene

### 2.1 Kildefiler

Kildefilene fra forlaget er vanligvis en av to typer. Ordbøkene er ofte tilgjengelig i form av trykksatsfiler, med andre ord filer som er beregnet på den maskinen som skal trykke ordboka. Vanligvis er slike filer en eneste lang sekvens av tegn og koder. Tegnene er naturligvis de tegnene som kommer fram i trykk, mens kodene er koder som bestemmer utseendet og posisjonene til tegnene, for eksempel koder for normal, uthevet eller kursiv skrift, koder for linjeskift, innrykk eller ny artikkel. Kodene kan også representere tegn som ikke finnes i det aktuelle tegnsettet, for eksempel matematiske tegn som integraltegn eller kvadratrottegn eller greske tegn som alfa og omega.

Figur 1 nedenfor viser et eksempel på hvordan en trykksatsfil kan se ut, og hvordan den korresponderende delen av ordboka ser ut. Eksempelet er hentet fra *Fremmedordboka*, og oppslagsordet er *abacis*t med litt omkringliggende tekst.

```
fra««<->Hamburg.<QL>««<UF2>abaci<ek66>st, <1>en, g
r., regnemester. - <3>a<ek66>bacus, <1>en, se««<->
abakus.<QL>««<UF2>abaddo<ek66>n, <1>hebr., eg. øde
```

```
ab, prep., lat., av, fra, f.eks. ab Hamburg, fra
Hamburg.
abaci'st, en, gr., regnemester. - a'bacus, en, se
abakus.
abaddo'n, hebr., eg. ødeleggelse, undergang, i
Det gamle testamente bet. for dødsriket.
```

Figur 1 - Eksempel på trykksatsfil og hvordan den ser ut i ordboka.

Som vi ser er nesten alle kodene avgrenset av mindre-enn og større-enn-tegn, "<" og ">". Vi ser at noen koder er koder for tekstattributter som uthevet, "<3>" eller kursiv, "<2>", mens andre koder representerer spesialtegn, som trykkaksenten ', "<ek66>". "Bomba" ("⌘") står her for tegn som ikke kan vises på trykk, i dette eksempelet tilsvarer to "bomber" et linjeskift.

Den andre typen kildefiler er på mange måter lik den første typen, men slike filer har en klarere struktur og ofte inneholder de opplysninger som ikke direkte kommer fram i den trykte utgaven av ordboka. Filene er gitt en klarere struktur ved for eksempel å ha linjeskift (i stedet for koder for linjeskift) der det er linjeskift i boka og gjerne litt "luft" mellom artiklene. Men den viktigste forskjellen fra trykksatsfiler er at de strukturte filene ofte inneholder feltkoder som ikke kommer fram i den trykte ordboka. Feltkodene er koder som forteller hvilken type opplysning som følger etter koden, for eksempel ordhistorie, definisjon eller uttrykk der oppslagsordet forekommer.

Figur 2 nedenfor viser et eksempel på hvordan en strukturert kildefil med feltkoder kan se ut, og hvordan den korresponderende delen av ordboka ser ut. Eksempelet er hentet fra *Bokmålsordboka*, og oppslagsordet er *fell* med litt omkringliggende tekst.

```

=
NB001      fell
NB001a     M1
TR007
..OPP      ##$Cfell@ m1
..ETY      (norr $Bfeldr@)
..DEF      $C1@ dyrepels
..DEF      $C2@ skinn til overbredsel
..UTR      $Bkrype under f-en@
=
NB001      fellah
NB001a     M1
TR007
..OPP      ##$Cfellah@ m1
..ETY      (fra arab 'bonde')
..DEF      egyptisk bonde

fell m1 (norr feldr) 1 dyrepels 2 skinn til
      overbredsel krype under f-en
fellah m1 (fra arab 'bonde') egyptisk bonde

```

Figur 2 - Eksempel på strukturert kildefil med feltkoder og utseendet i ordboka.

Vi ser at alle feltkodene finnes helt til venstre på hver linje. De fleste feltkodene er ganske intuitive, men en forklaring på noen av kodene følger. "NB001" er en kode som kommer foran selve oppslagsordet. Her er oppslagsordet uten eventuelle stammestreker eller trykkaksenter. Dersom flere oppslagsord deler samme artikkel, som for eksempel *felgbrems* og *felgbremse*, vil vi ha ett oppslagsord på en "NB001"-linje og det andre på en "NB002"-linje. Feltkoden "..ETY" kommer foran opplysninger om etymologi (ordhistorie), mens "..DEF" kommer foran selve definisjonene. Som vi ser, kommer ulike definisjoner på separate "..DEF"-linjer. Fra eksempelet kan vi dessuten se at det også for denne ordboka er koder for

skriftattributter, her "\$B" for kursiv og "\$C" for uthevet. Denne skrifttypen oppheves av en "@"-kode.

## 2.2 DAM-formatet

Dictionary Access Method (DAM) er et databasesystem spesielt beregnet for ordbøker. DAM lagrer ordbøkene i et format som også kalles DAM. Det som kjennetegner DAM-formatet, er at hver post i databasen kun består av to felter, et felt for selve oppslagsordet og et informasjonsfelt for informasjonen om oppslagsordet. Hva informasjonsfeltet inneholder, vil avhenge av hvilken type ordbok det er. En synonymordbok vil ha ulike synonymer til oppslagsordet, mens en tospråklig ordbok vil ha ulike oversettelser av oppslagsordet. I begge tilfeller kan informasjonsfeltet inneholde opplysninger om ordklasse, bøyning og lignende. Men uansett hvor mye eller lite som ligger i informasjonsposten, så vet ikke DAM-systemet noe om hva de forskjellige opplysningene er, siden alle opplysningene ligger i det samme informasjonsfeltet. Som navnet antyder arbeider DAM-systemet med dataene i databasen på samme måte som en vil bruke ei trykt ordbok, det vil si via oppslagsordet. DAM-systemet tilbyr alle operasjoner en kan forvente av et databasesystem, som innsetting, sletting og oppdatering av poster, samt ulike måter å hente informasjon fra databasen på, for eksempel ved hjelp av oppslagsord eller ordnummer. DAM tilbyr også mulighet for at brukeren selv kan bestemme sorteringsrekkefølgen for oppslagsordene i ei ordbok. Det betyr at DAM kan benyttes både for filer i ASCII og EBCDIC-format, og vi kan for en gangs skyld få våre "eksotiske" 3 tegn æ, ø og å sortert i riktig rekkefølge! I tillegg er det mulig å be om at blanke, bindestreker og lignende skal overses ved sortering, slik at *ad hoc* sorteres som *ad hoc*, eller for eksempel at tall skal sorteres som om de var skrevet med bokstaver, slik at *A4* sorteres som *Afire*.

## 2.3 LDB-formatet

En klar begrensning ved DAM-formatet er at opplysningene i informasjonsfeltet ikke er skilt fra hverandre. *Lexical Data Base* (LDB) er et ordboksformat som ikke har denne begrensningen. LDB har alle fordelene som DAM-formatet har og kan i tillegg skille de ulike opplysningene om oppslagsordet. Hver post i databasen har et vilkårlig antall felter hvor de ulike opplysningene plasseres, for eksempel felt for homografnummer, ordklasse og selve definisjonen. Feltene har en hierarkisk struktur som minner om en trestruktur. Denne strukturen svarer til brukerens oppfatning om organiseringen av de ulike opplysningene i en ordartikkel.

Figur 3 nedenfor viser hvordan en post i LDB-formatet ser ut (logisk). Oppslagsordet er *fell* fra *Bokmålsordboka*, og hvordan dette ser ut i den trykte utgaven av ordboka kan finnes i figur 2. Som vi ser, består treet av to typer greiner, terminerende og ikke-terminerende. Terminerende greiner representerer de feltene i databasen som inneholder de ulike opplysningene om et ord. "ordklasse" og "defnummer" er eksempler på slike greiner. De ikke-terminerende greinene er de som skaper den hierarkiske strukturen. Disse greinene kan ikke inneholde ordopplysninger, men inneholder i stedet andre greiner. "homografnode" og "defnode" er eksempler på ikke-terminerende greiner. Vi ser blant annet at opplysninger som tilhører samme definisjon, er plassert under samme "defnode". Vi ser også at en del opplysninger som ikke står eksplisitt i ordboka, er utledet og finnes i treet. Dette gjelder for

eksempel ordklassen substantiv, som her er utledet fra bøyningsskoden. Dessuten er den komprimerte formen *f-en* i uttrykket blitt utledet til *fellen*.

```

-entry
-homografnode
  -oppnode
    -oppslagsord : fell
    -skriveform : hovedform
    -representasjon : fell
    -ordklasse : substantiv
  -bøyingsnode
    -genus : m
    -bøyingskode : m1
  -ordhistnode
    -språk : norr
    -ord : felldr
  -defnode
    -defnummer : 1
    -definisjon : dyrepels
  -defnode
    -defnummer : 2
    -definisjon : skinn til overbredsel
  -uttrykksnode
    -uttrykk : krype under fellen

```

Figur 3 - Eksempel på den hierarkiske strukturen i en LDB-fil.

### 3 Konvertering til DAM-format

For å konvertere ei ordbok til DAM-format skal det altså lages et program som leser kildefilene, analyserer disse og legger de riktige opplysningene inn i databasen. Det finnes grensesnitt til DAM-systemet i flere ulike programmeringsspråk, blant annet Prolog og REXX, som er et Pascal-lignende språk. REXX er det språket som er benyttet i den framgangsmetoden som er beskrevet her. Programmets oppgave er å finne oppslagsordet med tilhørende informasjon fra kildefilene.

Hvor vanskelig denne oppgaven er, avhenger svært av hvordan kildefilene ser ut. For strukturerte kildefiler er det vanligvis ikke altfor vanskelig å kunne skille oppslagsord og informasjon. Kildefilene for *Bokmålsordboka*, jamfør figur 2, har oppslagsordene skrevet i korrekt skriveform på egne linjer ("NB"-linjer). Med korrekt skriveform menes her uten eventuelle stammestreker "l" og trykkaksenter "'" og alle tilder "~" er erstattet med den korrekte stammen. Det er derfor vanligvis en relativt enkel oppgave å finne oppslagsordet for slike kildefiler. Siden det kan være flere oppslagsord for hver artikkel, som for eksempel *felgbrems* og *felgbremsse*, må alle oppslagsordene tas vare på til hele ordartikkelen er lest inn. De neste linjene settes sammen for å lage selve informasjonsfeltet. Det kan lønne seg å ta vare på alle feltkoder og alle skrifttypekoder i disse linjene. Feltkodene er svært nyttige dersom ordboka skal konverteres til LDB-format. WordSmith har mulighet for å undertrykke disse slik at feltkodene ikke vises når brukeren slår opp et ord. Når hele ordartikkelen er lest, legges oppslagsordene med tilhørende informasjonsfelt inn i databasen. Det er naturligvis

mulig å bearbeide informasjonsposten før ordet legges inn i basen. Dette er aktuelt ved for eksempel erstatning av komprimerte former, som i *krype under f-en*, som kan utledes til *krype under fellen*. Etter at ordet er lagt inn i databasen, er det klart for å lese neste ord, inntil hele kildefilen er lest.

For trykksatsfiler kan oppgaven å skille oppslagsord og informasjon om ordet være litt vanskeligere. I dette tilfellet må programmet selv finne start og slutt på hver ordartikkel. Siden hvert avsnitt i ordboka kan inneholde flere artikler, jamfør *abacis*t og *abacus* i figur 1, må alt som står skrevet med uthevet skrifttype undersøkes. Dersom uthevet kun er benyttet for oppslagsordene, er dette relativt greit. Men dersom uthevet også er benyttet til andre ting, for eksempel definisjonsnummer eller inne i definisjonene, må programmet sjekke om det som står i uthevet er et nytt oppslagsord eller om det er en del av informasjonen om forrige oppslagsord. I enkelte tilfeller kan det være vanskelig for programmet å avgjøre dette, særlig i de tilfeller der uthevet er tillatt også inne i definisjonene. I tillegg til dette må programmet ta vare på stammer, slik at tilder "~" i seinere oppslagsord kan erstattes korrekt.

Figur 4 nedenfor viser et slikt eksempel. Tilden i *~brems(e)* må erstattes med den stammen som virker på dette stedet, *felg*. Dessuten må parentesene oppløses, slik at det blir to oppslagsord, *felgbrems* og *felgbremse*, med samme informasjon.

```
felg ml (fra lty) hjulring som slange og dekk
sitter rundt sykle på f-en / overf: være helt på
f-en utslitt, nedbrutt ~brems(e) på sykkel:
bremse på felgen, t forskj fra navbrems(e)
```

Figur 4 - Eksempel på ordartikkel med tilde og parenteser.

#### 4 Konvertering til LDB-format

Konverteringen av ei ordbok til LDB-format foregår på en helt annen måte enn til DAM-format. Konverteringen utføres ved hjelp av en analysator kalt Dictionary Entry Parser (DEP). En forutsetning for å bruke DEP er at ordboka er i DAM-format. Oppgaven nå er å analysere informasjonfeltet i DAM-postene for å finne de ulike opplysningene i dette feltet. Analysen foregår på grunnlag av et sett med grammatiske regler som sier hvordan opplysningene i ordboka er organisert (i forhold til hverandre). Ei liste over hvilke skilletegn eller "merker" (eng: token) som er brukt er også nødvendig. Denne lista inneholder alle tegnsekvenser som er benyttet i ordboka for å skille de ulike opplysningene fra hverandre. Lista vil derfor inneholde eventuelle feltkoder, koder for skrifttyper samt andre skilletegn som komma, semikolon og ulike typer parenteser. Disse skilletegnene gis symbolske navn.

Figur 5 nedenfor viser noen linjer fra lista benyttet for *Bokmålsordboka*. Vi kjenner igjen både skrifttypekoder og feltkoder fra tidligere eksempler. Det symbolske navnet lengst til høyre på hver linje kan velges fritt, men bør naturligvis ha ett fornuftig navn, slik som "font(normal)" for normal skrifttype.

De grammatiske reglene må ha en spesiell DEP-syntaks. Denne syntaksen ligger nær opptil Prolog-syntaks, med tillegg av en del spesialoperatorer. Figur 6 nedenfor viser noen grammatiske regler som finner definisjonsnummeret fra en definisjon i *Bokmålsordboka*.



```

delim("$A", font(normal)).
delim("$B", font(kursiv)).
delim("$C", font(uthevet)).
delim("@", fontend).

delim("^", linend).
delim("!!", entryend).

delim("TR007", entrybeg).
delim("..OPP", opp).
delim("..ETY", ety).
delim("..DEF", def).
delim("..UTR", utr).

```

Figur 5 - Eksempel på liste over skilletegn for DEP.

Definisjonsnummeret, "Defnr", er her sagt å være noe som ligger mellom skrifttypekoden for uthevet, "font(uthevet)", og koden som opphever denne skrifttypen, "fontend". Regelen "num\_string" passer dessuten på at definisjonsnummeret kun består av tall. Reglene "num\_string" og "defrec2" er ikke vist her, men må være definert et annet sted. "opt"-operatoren er en av DEPs tilleggsoperatører. Den tillater argumentet i parenteser å være "optional", det vil si frivillig. Her betyr det at i noen tilfeller har vi et definisjonsnummer (når antall definisjoner er to eller flere), i andre tilfeller mangler dette nummeret (når det kun er en definisjon).

```

defrec    ==>  -def :
                opt(defnummer) :
                defrec2.

defnummer ==>  -font(uthevet) :
                -Defnr :
                -fontend :
                num_string(Defnr).

```

Figur 6 - Eksempel på noen grammatiske regler for DEP.

Reglene virker på samme måte som Prolog-regler, enten så lykkes reglene, eller så mislykkes de. I kulissene blir de grammatiske reglene faktisk oversatt til Prolog-regler av analysatoren. Dersom de grammatiske reglene lykkes for et bestemt ord, betyr det at analysatoren har klart å analysere hele informasjonsfeltet og plassert opplysninger på greiner i treet i henhold til de grammatiske reglene. Dersom reglene mislykkes, kan dette ha flere årsaker. Den vanligste er at de grammatiske reglene ikke er omfattende nok til å ta hensyn til alle de ulike sammensetningene av koder og tekst. En annen årsak kan være at trestrukturen ikke er omfattende nok til å ta vare på alle ulike typer opplysninger i ordartikkelen. En tredje årsak som forekommer relativt sjelden, er direkte feil i kilden. Eksempel på slike feil kan være at uttrykk står i

normal skrift i stedet for kursiv, eller at forkortelsen for språket italiensk, "it.", mangler punktum ved en bestemt forekomst.

En av DEPs styrker er at de grammatiske reglene kan kombineres med vanlige Prolog-regler. Man kan derfor bruke vanlige Prolog-regler for å manipulere og kontrollere data før de legges inn i LDB-basen. Ofte er det nemlig slik at en tekststreng i en bestemt posisjon kan gi en av flere ulike typer opplysninger. Første opplysning etter selve oppslagsordet i Bokmålsordboka kan være en ordklasse, for eksempel "adj", eller en bøyingskode, for eksempel "m1". Disse to opplysningene skal inn på to forskjellige greiner, og reglene må derfor kontrollere at riktig opplysning legges inn på riktig plass.

Figur 7 nedenfor viser en helt vanlig Prolog-regel som avgjør om argumentet X er en av de kjente ordklassene. Denne regelen kan benyttes i de grammatiske reglene for å sikre korrekthet på opplysningene som legges inn i basen.

```
er_ordklasse(X)      <- Ordklasseliste=(
                    "adj"."adv"."art"."interj".
                    "konj"."pref"."prep"."pron".
                    "subst"."suff"."tallord"."v".
                    nil) &
                    member(X,Ordklasseliste).
```

Figur 7 - Eksempel på Prolog-regel som kan kombineres med grammatiske regler.

## 5 Avslutning

De fleste av dagens ordbøker er laget uten det formål for øyet at de i framtida skal kunne brukes i elektronisk form. Og dessverre er det noen av dem som vil lide litt av det når de forsøkes brukt i elektronisk form. Ei ordbok er som kjent et meget stort verk, og flere personer er vanligvis involvert i redigeringen. Det er derfor vanskelig å sørge for at hele ordboka er redigert på en konsekvent måte. Men viktigheten av dette kan ikke undervurderes dersom ordboka skal benyttes i elektronisk form. Mennesker har andre evner enn en datamaskin. I ei trykt ordbok kan et menneske forstå at en tekstsekvens er et uttrykk selv om den ikke står i kursiv skrift slik den burde ha gjort. Dette er vanskeligere å få en datamaskin til å forstå. Derfor er det veldig viktig at ordboka er redigert på en så konsekvent måte som over hodet mulig.

## Litteratur

- Berulfsen, Bjarne og Dag Gundersen. 1986. *Fremmedordbok*. 15. utg. Kunnskapsforlaget, Oslo.
- Byrd, Roy. 1989. *LQL User Notes - An Informal Guide to the Lexical Query Language*. IBM Technical Report, Lexical Systems Project, I.B.M. Thomas J. Watson Research Center, Yorktown Heights, New York.

- Byrd, Roy, Gustaf Neumann og Seved Andersson. 1988. *DAM - A Dictionary Access Method*. IBM Technical Report, Computer Science Department, I.B.M. Thomas J. Watson Research Center, Yorktown Heights, New York.
- Kirkeby, W.A. 1983. *Norsk-Engelsk Ordbok*. Oslo.
- Landrø, M. I., B. Wangensteen et al. 1986. *Bokmålsordboka*. Universitetsforlaget, Oslo.
- Neff, Mary. 1988. *Using DAM: An introduction to the Dictionary Access Method*. IBM Technical Report, Lexical Systems Project, I.B.M. Thomas J. Watson Research Center, Yorktown Heights, New York.
- Neff, Mary og Branimir Boguraev. 1991. *Dictionary Entry Parser Reference Manual*. IBM Technical Report, Lexical Systems Project, I.B.M. Thomas J. Watson Research Center, Yorktown Heights, New York.
- Neff, Mary og Branimir Boguraev. 1991. *From Machine-Readable Dictionaries to Lexical Data Bases*. IBM Technical Report, Lexical Systems Project, I.B.M. Thomas J. Watson Research Center, Yorktown Heights, New York.
- Neff, Mary og Roy Byrd. 1988. *WordSmith User's Guide: Version 2.0*. IBM Research Report RC13411, I.B.M. Thomas J. Watson Research Center, Yorktown Heights, New York.
- Neff, Mary, Roy Byrd og Omneya Rizk. 1988. *Creating and Querying Hierarchical Lexical Data Bases*. Proceedings of the Second ACL Conference on Applied Natural Language Processing. 84-92. Austin, Texas.

## "Vulgært" i tospråklige ordbøker

Termen vulgarisme betegner i grunnen tre ulike kategorier av leksemer: 1) Ord som etter sin form grovt avviker fra normen; 2) Ord som betegner "grove" begreper; 3) Ord som inneholder sterkt negative konnotasjoner. Tospråklige ordbøker bør ta med vulgarismer og forsyne dem med nøyaktige ekvivalenter med tilsvarende markører.

*Begrepet uanstendighet er fremmed for en språkforsker, og alle ord er like anstendige for ham.*

Ivan Baudouin de Courtenay  
(russisk lingvist, 1845 - 1929)

Markøren vulg. (vulgært), som brukes nokså ofte i ordbøker, betegner ulike karakteristika av et leksem. I foreliggende artikkel skal det gjøres et forsøk på å presisere dens innhold.

Begrepet vulgært i språket er umiddelbart knyttet til begrepet norm idet det vulgære defineres som det som ligger utenfor normen. Kriteriene for hva som er språknorm, varierer fra bok til bok, fra artikkel til artikkel, men stort sett nevnes følgende:

- 1) i hvilken grad fenomenet er utbredt;
- 2) om det er representert hos de mest autoritative forfattere
- 3) hvorvidt det svarer til vedkommende språks lovmessigheter.

Det understrekes i regelen at det ikke er nødvendig at alle disse tre kriteriene kan anvendes på et fenomen samtidig.

La oss i all korthet ta for oss disse kriteriene.

Et fenomenens utbredelse som sådant regnes ikke for å være noe pålitelig kriterium. Det er ikke sjelden språkfolk skriver omtrent sånn: riktignok er fenomenet utbredt, men bør fordømmes. (Det minner meg om et sitat fra den russiske forfatter Aleksander Kuprins novelle "Duellen": hele kompaniet går i utakt, bare løytnant Romasjov går i takt.) I grunnen dreier det seg om følgende: Siden ethvert språk forandrer seg, tenderer det hele tiden til å avvike fra den anerkjente norm (hvis den eksisterer), ellers opphører det å utvikle seg. Det store spørsmålet er hvilke novasjoner som er "riktige (gode, akseptable osv.)" og hvilke som er "gale (dårlige, uakseptable osv.)", det gjelder så å si å skille fårene fra bukkene. Med andre ord gjelder det å kunne forutsi hvilke former som har fremtiden for seg, og hvilke som ikke har det. Og det er språkfolk som regel ikke i stand til. Konstruksjonen *høres ut* (etter *se ut*) ble til å begynne med oppfattet som gal, senere som familiær (se *Norsk riksmålsordbok* (NRO) 1937), nå er den nøytral. På den annen side ligger vel konstruksjonen *føles ut* utenfor rammen av normen, men det er ikke godt å si om den ikke blir akseptert med tiden. I begynnelsen av vårt århundre ble uttalen /šl-/ i begynnelsen av et ord ansett for å være

vulgær, men nå er den kurant på Østlandet. Derimot er uttalen /'rušlan/ for Russland ikke standard østnorsk - foreløpig?

Man kan naturligvis alltid prøve å finne forklaringer på slike fenomener, f.eks. si at uttalen /rušlan/ er gal fordi /s/ og /l/ hører til forskjellige røtter, men forklaringen holder ikke stikk siden sandhi er typisk for norsk, jf. uttalen av navnet Bjørnson. Vi må simpelthen fastslå at slag og Bislett med /šl/ er ok, men Russland uttales i standard østnorsk med /sl/. Et eksempel til: Sekvensen /ŋš/ i norsk er nok sjelden, men den forekommer, jf. /br̥ŋšə/, /raŋ̃še:ra/, men uttalen /aŋšə/ av suffikset -anse eller /br̥ŋšə/ bronse regnes som vulgær.

Enda vanskeligere er det å finne noen objektive kriterier for hva som er riktig og hva som er galt når det gjelder det leksikalske nivå. Har en enhet vunnet hevd i språket, så er den riktig om den i og for seg er aldri så tvilsom. Uttrykket *i langsom kino* er gjengs, skjønt det ikke er logisk.

Det andre kriteriet, de autoritative forfatteres språkbruk, er heller ikke helt pålitelig. Bare det at slike forfattere som oftest hører fortiden til, gjør det mulig å forkaste deres språkbruk som foreldet. En liten liste over galt kjønn hos store forfattere er ganske lærerik: *en skred* (Bjørnson), *en hvelv* (Ibsen), *et tross* (Jonas Lie), *et lås* (Egge), *en smil* (Ansteins-son) osv. Det er velkjent at Ibsen stundom brukte *hang* der det skulle vært *hengte* og *brant* istedenfor *brente*. I grunnen kommer det an på om et eller annet eksempel fra en klassiker stemmer med det en moderne språknormerer anser for å være korrekt. Først da får eksemplet beviskraft (se, både Hamsun og Undset brukte denne formen). I motsatt fall kan man alltid si: Nei, siden Hamsun og Undsets tid har språket forandret seg betraktelig, og det som var riktig på deres tid, er nå foreldet.

Dette kan illustreres ved ordet *regn*. Ibsen og mange andre brukte ordet i direkte forstand som felleskjønn. I 1921 skrev Aug. Western: alltid en regn av granater (Western 1921). NRO noterer angående felleskjønnsformen: alm. i bet. 2 (noget som strømmer ned, daler, faller ned i stor mengde). Både *Riksmålsordboken* (Guttu 1977) og *Bokmålsordboka* (Landrø 1986) registrerer ordet som bare intetkjønn.

Det tredje kriteriet er enda vagere, særlig når det gjelder det leksikalske nivå. Hvordan kan man bevise at uttrykket *i langsom kino* svarer til eller tvertimot strider mot det norske språks lovmessigheter?

Konklusjonen av ovenstående resonnement er ganske banal. Normen for et språk er det som språknormerere anser for å være riktig, det som stemmer overens med deres språkfølelse. Alle de nevnte kriterier tas i bruk bare når de egner seg til å bekrefte normererens dom, ellers forkastes de. Og det er vel ikke noe galt i det, men litt subjektivitet er uunngåelig.

Etter denne innledning kan vi nå gå over til å drøfte begrepet vulgært i et språk. Hva vil det egentlig si?

NRO bruker markøren *vulg.* ikke så sjelden, men den defineres ikke i Innledningen. Selve ordet *vulgær* defineres i ordboken på følgende måte: *mest foreld.* (om sprog, sprogform) som brukes av de brede lag; folkelig.

Men dermed er problemet langt fra løst. NRO skjelner åpenbart mellom *vulgært* og *folkelig* (til tross for definisjonen som er anført ovenfor), samme ord forsynes ofte med disse to markørene samtidig, jf. *dritt* ... 1) vulg., folk. ... 2) vulg., folk. ...// vulg., folk. (ell. fam.) ... 3) vulg., folk. ell. fam. ... Denne rekkefølgen er ikke stående, jfr. *dere* ... 2) folk., vulg. til én (som en ikke er dus med, svarende til De og skrevet med stor forbokstav) ... . Markøren folk. kombineres i NRO også med andre markører, jfr. *å* pron. og *å* adv. som har følgende karakteristikker: dial., folk., bym.

Foruten disse to så å si materielle markører - vulg. og folk. - har vi en "usynlig" markør som brukes av leksikografer i mange land og som aldri omtales i eksplisitt form. Jeg mener utelattelse av et ord, det at det i det hele tatt ikke tas med i en ordbok. Dette gjelder i første rekke obskøne ord. Der slike ord tas med, forsynes de med markøren tabu e.l.

Ser vi nærmere på ordene som er forsynt med markørene vulg. eller folk. (eller i det hele tatt mangler i ordbøkene), så viser det seg at de stort sett danner tre klasser.

Den første klassen dannes av ord og ordformer som anses for å være folkelige eller vulgære etter sin form, mens deres innhold, deres betydning er nøytral.

Enhetens folkelige eller vulgære form kan være fonetisk eller grammatisk.

Enhetens folkelige eller vulgære fonetiske form kan være av segmental og/eller suprasegmental karakter.

Avvik fra normen når det gjelder de folkelige eller vulgære enheters segmentale sammensetning kan være regelmessige og uregelmessige. Regelmessige avvik skyldes den folkelige talens eller vulgærtalens lydige system, og behøver ikke (og kan ikke) registreres i ordbøker. Et eksempel på et slikt avvik kan være den østnorske måten å uttale sluttsekvensen *-el* på, dvs. med tykk l (nøkkel, Bibel), eller uttalen av suffikset *-anse* som /-aŋsə/. Ved uregelmessige avvik menes her et ords eller en ordforms segmentale sammensetning som det ikke kan stilles opp en generell regel for, f.eks. *å* 'hva', *å* 'hvor', *åfferlåffer* 'hvorfor', *åsså/også* 'å (infinitivsmerke)', /læ'rin:ə/ 'lærerinne' o.l.

Vulgære eller folkelige suprasegmentale former er stort sett regelmessige (systematiske). Den viktigste typen i norsk er trykk på første stavelse og tonem 2 i ord som *stasjon*, *bibl(i)otek*, *kultur*, *sørtere* o.l. Uregelmessige (usystematiske) avvik i trykk/tonem er vel sjeldne, jf. *vindue*.

Folkelige eller vulgære fonetiske avvik kombineres ikke så sjelden med grammatiske. De kan nok kalles for morfofonemiske vulgarismer. Dette kan illustreres ved rekken *vinni*, *finni*, *drekki* o.l., typen *stæle* - *stælte*.

Rent grammatiske vulgarismer (f.eks. syntaktiske) skal det ikke gås inn på her siden det er ord og ordformer som interesserer oss som leksikografer. Vi tar for oss de former (ord og deres varianter) som anses for å være folkelige eller vulgære. Særlig mange slike enheter er karakteristiske for pronomener: *Dere*, *vårs*, *våres*, *henner*, *vems*, 'a 'henne', *oss* 'vi'. Mindre folkelig, snarere familiær, er formen *dem* 'de', som i NRO betegnes som dial. ell. folk.

Den annen klasse omfatter betegnelser for "grove" begreper, dvs. ord som betegner det som i samfunnet anses for å være uanstendig, uestetisk osv. Det er tabubelagte ord, "primære" benevnelser på nevnte begreper. (Det at disse direkte benevnelser er primære, fremgår av at evfemismer som brukes istedenfor dem, er litterære - enten hjemlige eller latinske - ord.)

Det er klart at det eksisterer en tabuskala, at noen ord er mer tabubelagt enn andre. Det er også klart at bruken av tabuord er sosialt betinget og at den forandres med tiden.

Denne gruppen vulgarismer kan kalles for leksikalske.

Den tredje klasse utgjøres av ord med ulike sterkt negative vurderingskonnotasjoner. Formen av slike ord er ikke nødvendigvis folkelig eller vulgær, begrepene de betegner, kan være nøytrale, det er bare den sterkt negative konnotasjonen, grovheten, som gir slike ord folkelig eller vulgær karakter fordi de bare brukes på nevnte stilplan. I grunnen betegner ordene *ansikt* og *tryne* samme begrep, forskjellen er bare den at det siste ordet er sterkt nedsettende, foraktelig osv. Med andre ord består denne klassen av grove betegnelser for (mer eller mindre) nøytrale begreper.

De kan kalles for emotive vulgarismer.

Skillene mellom disse klassene er selvfølgelig flytende, det finnes gråsoner. F.eks. kan begrepet være nokså uestetisk, men ikke direkte tabubelagt, og én av betegnelse kan være nøytral uten å være en evfemisme, mens en annen kan være en vulgarisme (jfr. *neseslim - snørr*). Videre finnes det en gruppe ord som er leksikalske vulgarismer i direkte betydning (*dritt*) og emotive vulgarismer i overført betydning (*han er en stor dritt*). Særlig i kollokasjoner kan det kombineres trekk av alle disse klasser (*kjøss meg i rævva*).

Som det fremgår av det som er sagt ovenfor, kan disse klassene inndeles videre i underklasser. La oss for enkelthetens skyld inndeles hver slik klasse i to underklasser (skjønt enhetene i disse klassene danner kontinuerlige skalaer som det er vanskelig å kvantifisere). Da får nøytrale betegnelser indeksen 1, "svake" vulgarismer indeksen 2 og "sterke" - 3. På denne måten får vi 26 mulige kombinasjoner (kombinasjonen 1-1-1 er ingen vulgarisme, f.eks. *gran* som er korrekt etter formen, ikke betegner noe grovt begrep og ikke har noen negative konnotasjoner). Noen eksempler: 2-1-1 *dem* (istedenfor de), 3-1-1 /*konferanš*/ 1-2-2 *snørr*, 1-1-3 *tryne* (om menneskeansikt), 1-2-3 *dritt* (om menneske) osv. Evfemismer kan fremstilles som 1-3-1 eller 1-2-1. Noen kombinasjoner er nok mulige bare teoretisk.

Det kan altså utarbeides tre sett markører for tre selvstendige skalaer: 1) Riktighets-skala (ordets forhold til normen); 2) Tabuskala (begrepets grovhetsgrad); 3) Vurderingsskala (ordets grovhetsgrad). Dette gir et tredimensjonalt koordinatsystem for et ord. Det gjenstår ennå å utarbeide et metaspråk for disse skalaer.

Og til slutt noen generelle bemerkninger angående den leksikografiske siden av problemet.

Bør en tospråklig ordbok være normgivende (normativ), dvs. bare inneholde det som ligger innenfor rammen av normen? Mange leksikografer (og min) faste overbevisning er at den ikke bør være det. En tospråklig ordbok bør gjenspeile objektivt det som virkelig brukes i et språk, uansett om det svarer til normens krav eller ei. Den bør hjelpe leseren (brukeren) til å forstå alle slags tekster, og det er ikke leksikografers oppgave å bestemme hva brukeren behøver å vite og hva han ikke behøver. Som eksempel på riktig løsning på problemet kan det nevnes Cappelens store Engelsk-norsk ordbok.

Naturligvis bør sanne enheter forsynes med markører som skal gi brukeren en adekvat karakteristikk av ordet. Og her er det mange vanskeligheter. Det er ikke tilfeldig at noen former i NRO er forsynt med flere markører samtidig, f.eks. *vårs folk.*, *fam*.

Og så kommer vi til det siste problemet i forbindelse med vulgært - problemet ekvivalenter. Den generelle regelen for tospråklige ordbøker - at det overalt der det er mulig, bør gis ekvivalenter og ikke forklaringer, er aktuell også her. Brukeren bør få ekvivalenter av to grunner. For det første meddeler en riktig valgt ekvivalent all den informasjon om ordet som trenges, altså dets betydning, stilistiske valør m.m. For det annet bruker en oversetter eller en tolk som oftest nettopp ekvivalenter og ikke stilistiske synonymmer eller omskrivninger.

## Litteratur

- Cappelens store engelsk-norsk ordbok*. 1988. Oslo.  
 Guttu, Tor. 1977. *Riksmålsordboka*. Oslo  
 Landrø e al. 1986. *Bokmålsordboka*. Oslo.  
*Norsk riksmålsordbok I-III*. 1937. Oslo.  
 Western, A. 1921. *Norsk riksmålsgrammatikk*. Oslo.

## Finska för utlänningar - uppgifter om ordböjning i den finsk-norska ordboken

I dag finns det endast en ordbok (ett ficklexikon) mellan finska och norska. En ny finsk-norsk ordbok är under arbete. Den är i första hand avsedd för norska brukare. I artikeln beskrivs finskans komplicerade ordböjning samt olika möjligheter att ge upplysningar om den i en ordbok. Ordboksförfattarna har valt ett system med konventionella former av uppslagsord försedda med en indexsiffra som hänvisar till böjningstabeller + ordstammar för ord vilkas stammar avviker markant från uppslagsorden.

I ordböcker kan information om grammatik ges på många olika sätt. Det kan finnas direkta uppgifter t.ex. om ordklassstillhörighet, generella konstruktions uppgifter t.ex. om verbens rektioner eller exempel på användningen av ett ord som samtidigt innehåller implicit grammatik. Min uppsats har sina rötter i den finsk-norska ordboken som är under arbete. Jag kommer att behandla frågor som gäller grammatisk information med tanke på finskans ordböjning.

År 1978 gav det finska förlaget WSOY ut ett finsk-norsk / norsk-finsk ficklexikon. Författarna var Turid Farbrege och Aili Kämäräinen. Det var den första ordboken mellan de två språken och hittills den enda.

Ficklexikonet var avsett först och främst för finländska användare och detta faktum återspeglar sig i bokens struktur. Den norsk-finska delen är två gånger så stor som den finsk-norska delen, boken har 23 000 uppslagsord från norska till finska och 12 500 från finska till norska.

I den finsk-norska delen ges alla förklaringar på bokmål. Det finns

- 1 uppgifter om de norska ordens böjning t.ex. *aikoa* akte -et, ha (hadde, hatt) til hensikt
- 2 information om ordklass i vissa specialfall t.ex. *aasialainen* (s.) asiat -en, (a.) asiatisk -
- 3 en del tilläggsförklaringar om ordens betydelse eller användning. Dessa ges inom parentes och på finska. t.ex. *aihe* (syy) grunn en; (taulun) motiv -et; (kirjoituksen) emne -et

Uppgifterna om de norska ordens böjning förekommer också i den norsk-finska delen, också de sammansatta ordens böjning finns med. I den norsk-finska delen finns ca 200 nynorska ord med en hänvisning till motsvarande bokmålsord. Sist i boken finns en översikt över norsk ordböjning jämte en förteckning över några starka verb samt deras böjning i bokmål och i nynorska.

Nu förbereds en ny ordbok mellan de två språken, enligt planerna skall manuskriptet vara fullbordat hösten 1992. Det blir en enkelriktad ordbok från finska till norska med 400 - 450 sidor eller 25 000 uppslagsord. Den nya ordboken är tänkt som en passiv ordbok för



norska brukare, så uppgifter om norskans böjning kommer inte att finnas med. Rättskrivningsnormen för norskan är bokmål.

De flesta ordböcker som hittills har gjorts mellan finska och ett främmande språk är aktiva ordböcker för personer med finska som modersmål. Det finns några undantag såsom István Papps Finn-magyar szótár (Budapest, 1970), Kalju Pihels och Arno Pikamäes Soomeeesti sonaraamat (Tallinn 1965 och 1986) och en i Petroskoj i Sovjetunionen tryckt ordbok från finska till ryska som är avsedd för ryskatalande användare.

På norska finns det åtminstone Einar Haugens Norsk-engelsk ordbok där norska behandlas som ett främmande språk. Den är i första hand skriven för amerikanska studerande som skall lära sig norska.

Oftast är ordböckerna från finska till ett annat språk avsedda för finska brukare. Ordboksanvändaren slår upp i ordboken för att finna det främmande språkets motsvarighet till det finska ordet för att kunna bruka ordet antingen i skrift eller tal. För honom behövs det inga grammatikaliska uppgifter om finskan. Han vet hur t.ex. ordet *käsi* 'hand' böjs, att det är en substantiv osv. Den nya finsk-norska ordboken är tilltänkt en norsk brukare, som skall förstå finska med hjälp av en finsk-norsk ordbok. I en *aktiv norsk-finsk ordbok* för personer med norska som modersmål skall uppgifter om finskans böjning utan tvivel finnas med. Men hur mycket grammatikalisk information skall det finnas med i en *passiv ordbok från finska till norska*?

Bo Svensén säger i sin Handbok i lexikografi så här:

"För att rätt kunna tolka den franska (= ett främmande språk H.S.) man har framför sig kan det visserligen ibland vara viktigt att ha tillgång till böjningsuppgifter, men det säger sig självt att man oftare söker sådana uppgifter vid översättning till franska (= det främmande språket H.S.)" (Svensén, 1987:11)

Det här är säkert sant. Men ordboksförfattarna har inte alltid lämnat bort dylika uppgifter om källspråket. T.ex. Haugens ordbok är försedd med rikliga uppgifter om norskan trots att boken är en passiv ordbok för engelsktalande amerikaner. En passiv ordbok används oftast i en situation där ordboksanvändaren har en text som han vill läsa. Han söker efter ordens betydelse. Orden står redan i böjd form. Och när det gäller finska ord är det just de böjda formerna som ger ordboksanvändaren gråa hår.

Finskan är ett starkt agglutinerande och syntetiskt språk, medan norskan är analytisk som andra germanska språk. Ett finskt ord ger information även av syntaktisk art, medan motsvarande information på norska ges främst genom ordföljd, formord och partiklar, som är fristående från huvudordet. Vi kan jämföra meningar som "Heittäytyisinköhän seikkailuun?" och "Tenk om jeg skulle kaste meg inn i et eventyr?". Finskan klarar sig med två ord medan samma mening uttryckt på norska har tio ord.

I finskan kan ett nomen ha fem stammar till vilka ändelserna fogas. Alla finska ord har en sk. vokalstam. En del ord har två sorters vokalstammar starka och svaga. Finskans ord för 'hand' *käsi* har förutom nominativformen *käsi*- en stark vokalstam *käte*- och en svag vokalstam *käde*-. Här är det fråga om stadieväxling mellan t och d. Vid sidan av vokalstammen har några ord en konsonantstam, som slutar på en konsonant t.ex. *käsi* har konsonantstammen *kät*-. Som femte stam räknas pluralisstammen *käs*- av *käsi*. Som ett annat exempel kan vi ta ordet *vuosi* 'år' som också har både svag och stark vokalstam (*vuode*-/*vuote*-) och konsonantstam (*vuot*-). Verben kan ha tre sorters stammar. Som exempel på verb

kan vi ta ordet *tehdä* 'göra': svag vokalistam *tee-*, stark vokalistam *teke-* och konsonantstam *teh-*.

Enligt Fred Karlsson (1983:356-357) har finskans nominaler försiktigt räknat minst 2200 produktiva böjda och med ändelser försedda ordformer. Detta tal uppnås om vi räknar med två numerusformer, 13 olika kasusformer, 6 possessivsuffixer, 12 enklitiska partiklar eller deras förbindelser. Enligt Karlsson är det svårare att räkna ut hur många olika ordformer ett finskt verb kan ha. Ett verb kan ha 528 olika finita ordformer, 324 infinita ordformer och 11 000 participformer, sammanlagt ca 12 000 olika grammatikaliska ordformer. Det måste understrykas att det inte är fråga om avledda ord, de skulle göra talen flerdubbla. (T.ex ord av typen *köysi* 'rep' - *köyttää* 'binda, fastgöra med rep' - *köytättää* 'låta binda el. fastgöra med rep' / *alkaa* 'börja' - *alkaminen* 'början, start' / *varma* 'säker' - *varmuus* 'säkerhet').

Ett nomen kan ha 2200 och ett verb 12 000 olika former. Svårigheterna blir ännu större i och med att olika ord böjs på olika sätt. Också här uppvisar finskan höga tal. Nykysuomen sanakirja (Nufinsk ordbok), det äldre standardverket över finska språket, har 85 nominala böjningstyper och 45 olika böjningstyper för verb. Suomen kielen perussanakirja, ett nytt standardverk, vars första del utkom i fjol, har ett förenklat system: för nomen finns det 51 olika böjningsgrupper, för verb 27. Det har gjorts försök att minska antalet böjningstyper. Knut Cannelin kunde i sin bok *Finska språket* (1932) nöja sig med fem nominala böjningstyper. Kaino Laaksonen och Anneli Lieko har i *Suomenkielen äänne- ja muoto-oppi* (1988) fyra nominala och fyra verbala böjningsgrupper med underavdelningar. Ju mindre antalet böjningstyper blir desto oftare stöter vi på oregelbundenheter som måste förklaras på lexikonnivå.

Vidare kompliceras ordböjningen av finskans stadieväxling, som gör att två former av samma ord kan skilja sig mycket från varandra.

Dessa tal ger en uppfattning om hur invecklad finskans ordböjning är. Systemet kräver en utförlig behandling. Men i en ordbok med ett begränsat antal sidor är vi tvungna att göra kompromisser. Ju mera utrymme vi ger åt böjningsuppgifterna desto mindre blir antalet uppslagsord. Det går inte att redogöra för systemet i en ordbok. Som Bo Svensén (1987:68) säger, får man begränsa sig till sådant som är oregelbundet eller på annat sätt oförutsägbart. Svårigheten ligger i det att man måste bestämma vad som är oregelbundet. Skall en given form av ett ord betraktas som oregelbunden eller oförutsägbart när vi har ett språk där det är möjligt att böja nomen på ca 50 olika sätt?

I början är det svårt för en ordboksanvändare att finna det rätta finska uppslagsordet i en ordbok. Ju mer han kan finska desto lättare blir det, allteftersom hans kännedom om ordböjning ökar. Är det möjligt att göra en ordbok som skulle hjälpa användaren att hitta i ordboken ord som i sina böjda former avviker från uppslagsformen? Vilka möjligheter har ordboksförfattarna att underlätta användarnas arbete?

En talare med finska som modersmål kan producera olika ordformer, det är en del av hans kommunikativa kompetens. I den morfologiska och morfofonologiska beskrivningen av finskan utgår forskarna från att det finns grundformer, ur vilka de konkreta ordformerna härleds. *Nominativ i singularis* är den lexikaliska grundformen för ett nomen och det är ju också den form som finns i ordböcker som uppslagsform, t.ex. *takki, kaappi, tyttö, reikä*. Som uppslagsform i en ordbok har vi av verb *den första infinitiven*. Formen består av infinitivstammen och infinitivmärket (*liikku-a, hyppi-ä, hake-a*). I löpande text förekommer infinitivformen inte så ofta som personformerna. Psykologiskt sett är infinitivformen inte en grundform. Den kan inte heller vara utgångsformen för språkets beskrivning. *Den starka vokalistammen* (*liiku-, hypi-, hae-*) har valts som grundform för verb i finskans morfologiska

och morfofonologiska beskrivning. Den aktiva indikativformen i presens 3:e person i singularis (*liikkuu, hyppii, hakee*) är den mest frekventa formen i texter - den förekommer oftare än någon annan verbform. De minsta möjliga fullständiga satserna består av ett subjekt ofta i nominativ singularis, och ett predikat i aktiv indikativ presens t.ex. *Koira haukkuu.* ('Hunden skäller.').

Kunde ordbokstraditionen här förändras? Kunde man tänka sig att vi för verb i stället för infinitivformen skulle ha en annan uppslagsform t.ex. den starka vokalstammen eller en personform i presens indikativ aktivum? Också nu finns det variation i val av uppslagsformer, alla ordböcker på olika språk har ingalunda nominativ resp. infinitiv som uppslagsform. I ordböcker med latin som källspråk är uppslagsformen för verb presens indikativ aktivum 1:a pers. i sg. I ordböcker från estniskan, för estlänningar, står verben i infinitiv III.

I Finland har det kommit ut några läroböcker för utlänningar som studerar finska, där ordstammarna är genomgående med i ordförteckningen (Hämäläinen-Lehikoinen-Steadman, 1990). I ordlistorna finns som uppslagsord också böjda former med hänvisning till grundformen som annars kan vara svår att hitta. Systemet underlättar elevernas arbete när de skall producera finska, men stammarna har också en viktig funktion för språkets förståelse. Stammarna kan ge eleverna stöd när de söker efter ordens betydelse. Finns det i texten ett ord som *kätensä* t.ex. i *nostaa kätensä* 'lyfta sin hand' kan användaren, om han hittar uppslagsordet *käsi* och om det finns uppgifter om stammarna i ordboken, få stöd för sitt antagande att *kätensä* är en annan form av ordet *käsi*.

*Olika möjligheter att ange ordens böjning:*

(A) Böjda former eller ordstammar som uppslagsord

t.ex. *käden* (gen.ent) se \**käsi* /*käde-* se \**käsi*  
*kätenä* (ess.ent) / *käte-* se \**käsi*  
*kättä* (part.ent) / *kät-* se \**käsi*

Dessa är de bästa alternativen för ordboksanvändaren. Tyvärr kräver de så mycket utrymme att de inte är användbara i praktiken.

(B) Konventionella uppslagsformer (nominativ sg./infinitiv) +

(a) en indexsiffra som hänvisar till böjningstabeller som finns först/sist i boken

t.ex. *käsi*<sup>27</sup> 'hånd'  
*kuunnella*<sup>67\*</sup> 'lytte, høre på'

\* visar att ordet har en stadieväxlande konsonant, bokstaven anger typen av stadieväxling (här nn:nt *kuunnella/kuuntelen*).

(b) ordstammar efter uppslagsordet

t.ex. *käsi* (*käde-/käte-/kät-*) 'hånd'  
*omena* (*omena-*) 'eple'  
*lapsuus* (*lapsuude-/lapsuute-/lapsuut-*)  
*lapsuus* (*-suude-/suute-/suut-*)  
*lapsuuls* (*-de-/te-/t-*) 'barndom'

Med det här systemet kommer vi att få med en stor mängd ord som har endast en stam som är likadan som själva uppslagsordet. Det är inte nödvändigt att ta med dylika fall.

En annan fråga är om stammarna skall ges i klartext eller som förkortningar. Vi har valt att skriva ut hela stammen då den är angiven.

## (c) ordstammar efter uppslagsordet för de frekventaste orden

Detta är ett alternativ som bygger på tanken att en nybörjare använder endast ett begränsat antal ord, som hör till de vanligaste i språket. Dessa skulle få markering av stammarna. Efter en tid lär eleven sig ordböjning och hittar de mindre frekventa orden även om inte stammarna finns med. Nackdelen med systemet är att ord som *omenta* och *lapsuus* som inte är svåra att hitta i en ordbok skulle förekomma med stammen.

## (d) ordstammar för de ord vilkas stammar avviker markant från uppslagsformen

t.ex. *apu* (avu-/apu-) 'hjälp' *ydin* (ytime-/y $\ddot{a}$ in-) 'kjerne'

*antaa* (anna-/anta-) 'gi'

*aalto* (aallo-/aalto-) 'bølge'

Ju längre inne i ett ord uppslagsformen skiljer sig från stammarna desto lättare hittar ordboksanvändaren det rätta uppslagsordet. Böjda former av ord där skillnad förekommer först i den 5:e bokstaven som i *kielto* (kiello-/kielto-) 'förbud' eller *keväät* (kevää-/keväät-) 'vår' eller ännu längre inne i ordet torde inte vara omöjliga att hitta i en ordbok.

Den nya finsk-norska ordboken skall läggas upp enligt alternativ (B) så att punkterna (a) och (b) kombineras, t.ex. *apu* <sup>1\*E</sup> (avu-/apu-) 'hjälp' eller *antaa* <sup>56\*J</sup> (anna-/anta-) 'gi'.

Vi är medvetna om att det är en kompromisslösning. Det begränsade sidoantalet gör att vi inte kan göra boken så brukarvänlig som vi egentligen skulle ha lust till.

Det finns många frågor utan svar, både sådana som har med språkinläring att göra och sådana som hör ihop med användning av ordböcker. Kunde vi besvara t.ex. följande frågor, skulle vi ha en tryggare utgångspunkt för vårt arbete.

- När börjar eleverna behärska finskans ordstammar? Hur länge behöver de hjälp?
- Behöver de i första början alls en ordbok eller klarar de sig med läroböcker och deras ordförteckningar?
- När går eleverna över till enspråkiga ordböcker som *Nykysuomen sanakirja* - om de överhuvudtaget gör det?

## Litteratur

- Farbregd, Turid - Kämäräinen, Aili. 1978. *Suomi - norja - suomi taskusanakirja*. Juva.
- Hämäläinen, Eila - Lehikoinen, Laila - Steadman, David. 1990. *Aletaan! Jatketaan!* (Sanasto - Vocabulary). Helsinki.
- Karlsson, Fred. 1983. *Suomen kielen äänne- ja muotorakenne*. Juva.
- Nykysuomen sanakirja*. 1976. Porvoo.
- Suomen kielen perussanakirja*. Ensimmäinen osa A-K. 1990. Helsinki.
- Svensén, Bo. 1987. *Handbok i lexikografi*. Stockholm.

Rik Schutz

## Redigeringsprogrammet EDI

Programvaren EDI er spesielt laget for redigering av ordbøker. Jeg vil ta utgangspunkt i Van Dales svensk-nederlandske ordbok for å forklare hvordan programmet arbeider. Deretter vil jeg sammenligne ordboksartiklenes struktur og innhold med den redaktørene tidligere måtte arbeide med. Til slutt vil jeg vise hvordan en ferdig ordboksartikkel ser ut når redaktøren arbeider med EDI.

### Tradisjonell ordboksredigering

Til tross for de mange fordelene med den tradisjonelle redigeringsmetoden, er det store ulemper med den. Det største problemet er systemets kompleksitet og den uforholdsmessig store oppmerksomheten en redaktør/oversetter må vie nokså tilfeldige overenskomster. For eksempel kan punktum eller ikke punktum etter en forkortelse, eller bruk av vanlig parentes i stedet for hakeparentes, gi alvorlige feil. En redaktør som må konsentrere seg om slike ting, kan i mindre grad konsentrere seg om det leksikografiske oversettelsesarbeidet. Det tar svært lang tid å lære konvensjonene. En annen ulempe er risikoen for at tekststykker utilsiktet blir ødelagt eller helt mistes på grunn av uerfarenhet eller skjødesløshet fra medarbeiderens side.

Et problem som fremfor alt forleggeren føler, er tidstapet som etterkontrollen medfører. Tidsmessig går det mer enn et halvt år mellom fullføringen av manuskriptet og den endelige utgivelsen av boken. Mer enn halvparten av denne tiden går med til kontroll og korrigerings av manuskriptets struktur.

### EDI

Det er derfor vi har forsøkt å utarbeide en arbeidsmåte som ikke har disse ulempene. Redaktøren må kunne konsentrere seg om sitt egentlige arbeid. Kunnskap om de forkortete navnene på byggestenene og de nøyaktige reglene for rekkefølgen av dem, ligger i programmet. Programmet inviterer til å arbeide i riktig rekkefølge og gir beskjed hvis noe galt hender.

Programmet er i stand til å lese tekster som har den tradisjonelle Van Dale-strukturen, slik at disse kan behandles med EDI. EDIs lagringsstruktur har karakter av en relasjonell database. En direkte kopling mellom EDI og fotosettemaskinen er for øvrig tenkelig, men den brukes ikke ennå.

EDI fastlegger dessuten av hvem og til hvilket tidspunkt det er arbeidet med en fil, noe som har klare administrative fordeler. Det gjør det også enklere å gi feedback.

## Oppbygningen av en artikkel

Vi skiller mellom oppslagsord på grunnlag av deres form, altså på grunnlag av staving. Dette kaller vi det høyeste eller det første nivå. Ord som har felles form, såkalte homografer, kan i flere henseender skille seg fra hverandre. Avhengig av disse forskjellene velges en bestemt typografisk gjengivelse i ordboken. Denne gjengivelsen er nokså tilfeldig. Vi gjør dette slik:

- Hvis ordklassen (grammatisk kategori) er forskjellig, blir homografer trykt med nummer, slik som for eksempel hetta som henholdsvis substantiv og verb (jf figur 1).
- Hvis den grammatisk subkategorien er forskjellig (kjønn ved substantiv, eller transitive versus intransitive verb), blir en inndeling med romertall brukt, slik som for eksempel ved hetta som verb.

Det at Esseltes svensk-engelsk ordbok, som vår baserer seg på med hensyn til den svenske delen, gjør dette helt motsatt, viser hvor vilkårlig gjengivelsen av denne type forhold er. For EDI er det nivået som ordklasseinformasjonen legges på (såvel kategori som subkategori) viktig; vi kaller det for det andre nivå, eller for lemma-nivå.

På dette nivået kan det også angis markeringer av typen fagterm-etiketter, anvisninger av stilregister og morfologiske særegenheter som for eksempel verbbøyning.

Det tredje nivå er nivået der betydningsforskjeller markeres. Jeg vil her ikke gå inn på de spesielle problemene som oppstår i den forbindelse.

For hver betydning kan det gis eksempler på kontekster som oppslagsordet forekommer i. Den vanlige termen for dette er eksempel. Ofte dreier det seg om kollokasjoner og om idiomer. Dette kaller vi det fjerde nivå. I de fleste av Van Dales ordbøker er det fjerde nivå også grafisk skilt fra det tredje. Alle nummererte betydninger av oppslagsordet følger umiddelbart etter hverandre, og først deretter, etter et iøynefallende tegn (en ruter), kommer eksemplene. Men dette systemet er ikke relevant for EDI, og jeg går derfor ikke nærmere inn på det her.

## Byggestener

På hvert av disse fire nivåene forekommer i nesten alle artiklene tilbakevendende serier av byggestener. Hver av disse byggestenene har et navn forkortet til fem bokstaver. For en nederlender er denne forkortelsen tilstrekkelig til at man kan avlede ordets betydning av den.

Eksempler:

TREFW	= TREFWOORD	= oppslagsord
STILA	= STIJLLABEL	= stilistisk etikett
HVERT	= HOOFDVERTALING	= hovedoversettelse

Det er fastlagt i detaljerte skjemaer på hvilke måter disse byggestenene kan sammenføres til en artikkel (se figur 2). Av dette skjema kan man se at bestemte elementer er obligatoriske (de liggende blokkene), mens andre er valgfrie (de stående blokkene med stiplet linje), og at rekkefølgen elementene forekommer i, er fastlagt. Det er også angitt hvilke elementer som kan forekomme mer enn en gang.

Den leksikalske informasjonens lagringsform er en enkel, sekvensiell ASCII-tekst. På basis av rekkefølgen av og navnene på byggestenene angir settedatamaskinen de korrekte typografiske verdiene. Oppslagsord blir automatisk uthevet, og tegn som komma og parentes plasseres automatisk. Før denne transformasjonen kan finne sted i fotosettemaskinen, må manuskriptet være 100% rent. Kontrollen av at det formelle i teksten er helt riktig (for eksempel om betydningene står i riktig rekkefølge, og om det er gitt eksempler på ikke eksisterende betydninger), skjer med spesielle kontrollprogram. Det første lemmaet hetta ser slik ut i denne strukturen:

```

<TREFW.>  hetta
<GRAMT.>  m/v
<GRMAZ.>  ~n
<NULCO.>  0.1
<RESUM.>  sterk v\ "arme
<HVERT.>  hitte
<NULCO.>  0.2
<LABEL.>  fig.
<HVERT.>  vuur
<HVERT.>  enthousiasme
<CIJCO.>  1.1
<ILLUS.>  @solens ~
<VVERT.>  de hitte van de zon
<OMKWD.>  hitte
<CIJCO.>  1.2
<VASTV.>  i @stridens ~
<VVERT.>  in het vuur van de strijd
<OMKWD.>  vuur

```

Den skisserte arbeidsmåten har flere fordeler. Lagringen i tekstform gjør desentralisert bearbeiding enklere. Mer enn førti redaktører fra tre forskjellige land har samtidig arbeidet med en spansk-nederlandsk ordbok, hver på sin PC. Kontrollen etterpå skjedde sentralt.

### Virkemåte

De fire nivåene i artikkelen som jeg nettopp nevnte, ser i EDI ut som i figur 3. Oppe til venstre er det første nivå gjengitt, altså listen over oppslagsord. Deretter følger lemma-nivået, altså de grammatikalske kategoriene. Til venstre, i andre blokk ovenfra, står betydningene med oversettelsene i blokken ved siden av. Under disse til venstre står eksemplene, og til høyre ved siden av oversettelsene.

Når vi starter EDI, ser vi det første ordet fra filen vi arbeider i. Det er med andre ord A hvis vi behandler hele ordboken under ett, eller et annet oppslagsord hvis vi arbeider med mindre omfangsrike filer. I den svensk-nederlandske ordboken disponerer vi nå over hele bokstaven H, og alle eksemplene i denne sammenhengen er dermed tatt derfra. Vi kan søke etter ordet vi vil se på to måter: gjennomgå listen med piltastene eller med søkefunksjonen. Det siste virker absolutt raskere. Hvis vi nå har søkt oppslagsordet hem, ser vi i blokken oppe til venstre en del av oppslagsordlisten som vårt oppslagsord er markert i. I figuren er markeringen gjengitt med understrekning av det valgte elementet. Vi ser at det finnes to

grammatikalske kategorier av ordet vi har søkt fram. Adverbet, på samme måte markert med understreking, har tydeligvis bare en betydning. Til betydningen hører en oversettelse og de fem eksemplene fra blokken under til venstre.

Redigeringen av en artikkel foregår på følgende måte: Redaktøren velger nivået han vil arbeide på og flytter markøren dit. Dette kan enten gjøres ved å gå til ønsket nivå (blokk) med piltastene, eller ved å taste inn tallet som svarer til blokken. Ved å taste inn 5 går markøren rett til den femte blokken, altså den med setningseksempler. Informasjonen i en slik blokk kan forandres ved at man taster inn bokstaven E for EDIT, eller ved å trykke på ENTER-knappen.

### Tilføying av eksempler

La oss gå ut fra at vi vil føye et setningseksempel til den første betydningen for hem som substantiv: nemlig "ett andra hem". Vi sørger for at markøren befinner seg ved den første kategori (substantiv) i den andre blokken (grammatisk kategori) og ved den første betydningen i den tredje blokken (betydninger). I feltet nedenfor vises de tidligere redigerte eksemplene for denne betydningen. Ved å gå til dette feltet og taste inn T for Tilføye, kommer det fram en tom redigerings skjerm nederst, som har skjema-layout. Samtidig vises alle ulike former for informasjon som kan forekomme her. Der skriver jeg eksemplet (se figur 4). Hvis et bestemt felt ikke er aktuelt, for eksempel fordi oppslagsordets ordklasse utelukker det, angir EDI dette. I de feltene hvor det kun er tillatt å fylle ut ett element fra en avgrenset liste, en forkortelse eller en standardisert kode som angivelse av flertall, viser EDI de tillatte mulighetene nederst, så snart markøren befinner seg i det aktuelle feltet.

Jeg kan for eksempel velge mellom fire slags eksempler: 1 idiom, 2 illustrerende eksempler, 3 kollokasjoner (Vaste Verbindingen på Nederlandsk) og 4 grammatikalske kollokasjoner. Etter at jeg har valgt en av dem, kan jeg velge riktige eksempler ved å oppgi tallet. Det er selvsagt også mulig å taste inn den riktige forkortelsen. Hvis jeg taster inn en markering som ikke finnes, nekter EDI å forlate dette feltet. Etter at hele skjemaet er fylt ut, kan jeg gå tilbake til et av de andre nivåene.

Hvis jeg ønsker en oversikt over hele artikkelen, kan EDI vise lemmaet i ordbok-layout ved hjelp av ett enkelt tastetrykk (V for Vis). Da får jeg en oversikt over alle betydninger og alle eksempler, sammen med de tilhørende oversettelsene.

EDI kontrollerer inndata ved forskjellige handlinger. I bestemte felt aksepterer programmet kun det foreskrevne innholdet, slik som for eksempel ved ordklasser og markeringer. Når et skjema er utfylt, kontrollerer EDI om innholdet er korrekt. Og når artikkelen forlates, altså ved neste oppslagsord, sjekker EDI hele artikkelens struktur.

Etter redaksjonens ønske er det mulig å velge å forlate en artikkel før den er perfekt. EDI kan gjenkjenne og registrere tyve forskjellige slags feil. Når redaktøren forlater en artikkel spør EDI om han er klar over at det registreres en feil, for eksempel at det ikke er fylt ut noen oversettelse. Hvis redaktøren svarer bekræftende, skriver EDI en kode foran oppslagsordet i oppslagsordlisten, som angir den alvorligste feilen. Etter at feilen er løst av redaksjonen, blir feilmeldingen erstattet av en melding om den neste store feilen. Først etter at hele lemmaet er strukturelt korrekt, vises oppslagsordet uten koder på skjermen. Det er selvfølgelig mulig å lete etter en spesiell type feil, slik at en redaktør kan konsentrere seg om å løse liknende redaksjonelle problemer.



## Man ser hva man gjør

Etter min mening er det på en måte tatt et skritt tilbake, men ikke i negativ mening; det er et skritt bakover som likevel fører til fremgang. I urtiden, for ca. ti, femten år siden, arbeidet redaktørene med blanke papirark når de skulle lage en ny ordbok, eller med forstørrede kopier av en eksisterende ordbok. Foran seg kunne de se nøyaktig hvordan manuskriptet ville se ut i satsen, og de skrev den ønskede eller forbedrede teksten, med alle typografiske særegenheter.

Forlagenes ønske om å kunne behandle manuskriptet på datamaskin, har i løpet av siste tiår ført til at våre redaktører har måttet arbeide med et manuskript som ikke lenger så ut som selve ordboken, og som omfattet mange forstyrrende elementer. Diakritiske tegn, slik som ä i värme, måtte koderes og de typografiske særegenhetene ble eliminert i manuskriptet. Etter at manuskriptet var ferdig, kunne forleggeren nyttiggjøre seg kontrollmulighetene som en datamaskin har, og dermed kunne satsen lages billigere enn før; ved å bruke en fotosettemaskin i stedet for å sette i bly. Siden manuskriptet kun inneholdt koder og ikke typografisk informasjon, kunne typografien, og dermed også omfanget, fastslås først i siste øyeblikk.

Tilbakegangen med EDI er at nå kan redaktørene på nytt arbeide på den tradisjonelle måten med en artikkel som viser verken mer eller mindre enn det de fører inn i den. De diakritiske tegnene gjengis slik de skal være, og kodene er ikke lenger synlige. Fremskrittet er at forleggeren fremdeles kan nyttiggjøre seg alle fordelene med et maskinleselig manuskript og i tillegg har garanti for at strukturen i manuskript er riktig.

## Sluttord

EDI har mange funksjoner og finesser for praktisk ordboksarbeid, og jeg håper med disse eksemplene å ha videreført noe av den entusiasmen som vi og våre redaktører har for dette verktøyet.

Til slutt vil jeg kort nevne de gunstige egenskapene ved programmet:

- \* Leksikografen trenger ikke å ha detaljkunnskaper om det valgte systemet.
- \* Leksikografen trenger ikke å huske symboler eller koder; mulighetene vises stadig i en hjelpeskjerm.
- \* Redaktøren ser de spesielle tegnene slik som de ser ut i satsen, selv om de blir lagret med en koding som utelukkende arbeider med standard ASCII-tegn.
- \* Programmet aksepterer ikke strukturelle feil.
- \* EDI aksepterer imidlertid et manuskript som ikke er komplett, men registrerer hva som er feil, slik at det på grunnlag av meldingene raskt kan utføres forbedringer eller tilføyelser på et senere tidspunkt.
- \* Programmet muliggjør svært hurtig søking i andre artikler, og fremmer dermed redaksjonell uniformitet i semantisk sammenhengende artikler som ikke følger hverandre alfabetisk.
- \* EDI gjør det enkelt å arbeide tematisk. Det er svært enkelt å søke på kjennetegn som ordklasse, en bestemt type markering, om det i det hele tatt forekommer visse elementer (en bestemt kode for fagområde) i en artikkel.

- \* En satssimulator kan til enhver tid gi redaktøren en oversikt over artikkelen i samme layout som i den endelige ordboken. Denne satssimulatoren er i stand til å vise den endelige bokens totale sideomfang. Når det gjelder et nytt produkt, er det svært viktig å vite omfanget på den ferdige delen av manuskriptet.

### Litteratur

- Stora svensk-engelska ordboken*. 1988. V. Petti and I. Rider (eds). Solna.  
*Van Dale handwoordenboek nederländsk-svensk/Nederlands-Zweeds*. Under utgivelse. G.L.M. Laureys & J. de Groot et al. Utrecht.  
*Van Dale handwoordenboek svensk-nederländsk/Zweeds-Nederlands*. Under utgivelse. G.L.M. Laureys & J. de Groot et al. Utrecht  
*Van Dale handwoordenboek Nederlands-Spaans*. 1992. P.J. Slagter et al. Utrecht.  
*Van Dale handwoordenboek Spaans-Nederlands*. 1992. P.J. Slagter et al. Utrecht

Ole Norling-Christensen

## Struktureret redigering af ordbøger

Det standardiserede strukturbeskrivelsessprog SGML gør det muligt at beskrive vilkårligt komplicerede tekststrukturer og at opmærke teksten i overensstemmelse med beskrivelsen. Herved adskilles præsentation og indhold, og teksten bliver velegnet til alle slags datamatisk behandling, herunder syntakskontrol, komplicerede søgninger, automatiske ændringer; genbrug af definerede dele af en ordbogs oplysninger, præsentation af samme tekst med forskelligt lay-out. Teknikken og nogle praktiske anvendelser beskrives med en dansk-engelsk ordbogsartikel som gennemgående eksempel.

En del af meningen med de synspunkter og den teknik, som her skal introduceres, er at flytte fokus for leksikografens opmærksomhed fra formen henimod indholdet. Under det daglige ordbogsarbejde skal der ikke længere tænkes i typografiske kategorier, men i indholdsmæssige. En gang for alle (omend med mulighed for senere ændringer) har leksikografen beskrevet væsentlige dele af sine redaktionsregler i et formelt sprog, gjort rede for, hvilke oplysningstyper ordbogen skal indeholde og for deres indbyrdes sammenhæng. Da de indholdsmæssige kategorier således på forhånd er fastlagt, kan leksikografen herefter koncentrere sig om, hvilke af kategorierne der skal tages i brug i hvert enkelt tilfælde, og hvad der skal puttes i dem. Denne målsætning kan opfyldes på lige så mange måder, som der findes databasesytemer; men hvis eksisterende ordbøgers (makro- og mikrostruktur skal indpasses i et databasesystem, når man oftest til så komplicerede modeller, at det nærmer sig det prohibitive. Lettere er det, når man starter forfra med et helt nyt ordbogsprojekt.

Vilkårligt komplicerede strukturer kan imidlertid beskrives i SGML, Standard Generalized Markup Language, og SGML kan bruges til en eksplicit strukturmarkering af ordbogsteksten. En ny tekst kan indskrives, og en eksisterende tekst rettes til, med en særlig SGML-editor, som understøtter strukturmarkeringen. Et syntakskontrolprogram, SGML-parseren, kan kontrollere om markeringerne i teksten er i overensstemmelse med strukturbeskrivelsen. Den strukturmarkerede tekst er velegnet til alle slags datamatisk behandling, herunder: indlæggelse i databaser; komplicerede søgninger; automatiske ændringer; udtrækning af definerede dele af ordbogens oplysninger til brug fx i andre ordbøger; præsentation ikke blot i trykt form, men også fx på dataskærm. På basis af DANLEX-gruppens undersøgelser og anbefalinger (DANLEX 1987) har softwarehuset TEXTware A/S og forlaget Gyldendal i fællesskab udviklet et redigeringssystem, GestorLEX, til ordbøger. Det kombinerer SGML-editor, -parser, og en lang række andre faciliteter.

Strukturbeskrivelsessproget SGML er en international standard (ISO 8879, 1986) beregnet til at beskrive, og styre behandlingen af, alle slags "dokumenter", fx forretningsbreve, EF-cirkulærer, ganske almindelige bøger, og altså også ordbøger. Dokumenterne beskrives som en træstruktur af *elementer*; et element kan bestå af andre elementer og/eller stumper af egentlig tekst, som er træets blade. Dokumentet er selv et element, træets rod.

### Indhold, struktur, præsentation

En grundlæggende tanke bag SGML er at adskille præsentation og indhold. Man beskriver, hvilken *slags* oplysninger dokumentet består af, og hvordan de hænger sammen, altså en *generisk* (arts-mæssig) beskrivelse og kodning. Man beskriver derimod ikke, hvordan disse oplysninger skal præsenteres på tryk eller fx på en skærm. Præsentationsoplysningerne (skrifter, skriftstørrelser, interpunktionstegn og andre separatorer, spring til ny side eller nyt afsnit, osv.) kan føjes til bagefter som en funktion af strukturen. Til forskellig brug af samme dokument kan dettes strukturoplysninger meget vel oversættes til forskellige sæt af præsentationsoplysninger, uden at der overhovedet skal ændres i selve dokumentet. Et eksempel herpå er de danske Gyldendals Elektroniske Ordbøger og de tilsvarende norske fra Kunnskapsforlaget og svenske fra Almqvist & Wiksell (tidl.: Esselte Studium): det til grund liggende dokument er det samme; men skærbilledet, fx Axelsen (1990), og bogsiden, fx Axelsen (1984), ser meget forskellige ud.

### Dokumenttypedefinitionen, DTD

Et fuldstændigt SGML-dokument består af en dokumenttypedefinition, DTD, som angiver reglerne for dokumentets struktur, samt selve dokumentet, hvori start og slut på hvert enkelt element er markeret. Et simpelt eksempel på en DTD er den generelle beskrivelse i *figur 1* af en (fag)bog; den er lånt fra (FORMEX 1985), men ændret en smule.

---

<!ELEMENT BK	(IP, MP, EP) >
<!ELEMENT IP	(TK, CN, PR) >
<!ELEMENT TK	(#PCDATA) >
<!ELEMENT CN	(HC, T+) >
<!ELEMENT (HC, T)	(#PCDATA) >
<!ELEMENT PR	(TP, P+) >
<!ELEMENT (TP, P)	(#PCDATA) >
<!ELEMENT MP	(CH+) >
<!ELEMENT CH	(TC, P+, F*, BI?) >
<!ELEMENT (TC, F)	(#PCDATA) >
<!ELEMENT BI	(TB, I+) >
<!ELEMENT (TB, I)	(#PCDATA) >
<!ELEMENT EP	(PS, G, IX) >
<!ELEMENT PS	(TS, P+) >
<!ELEMENT TS	(#PCDATA) >
<!ELEMENT G	(TG, I+) >
<!ELEMENT TG	(#PCDATA) >
<!ELEMENT IX	(TI, I+) >
<!ELEMENT TI	(#PCDATA) >

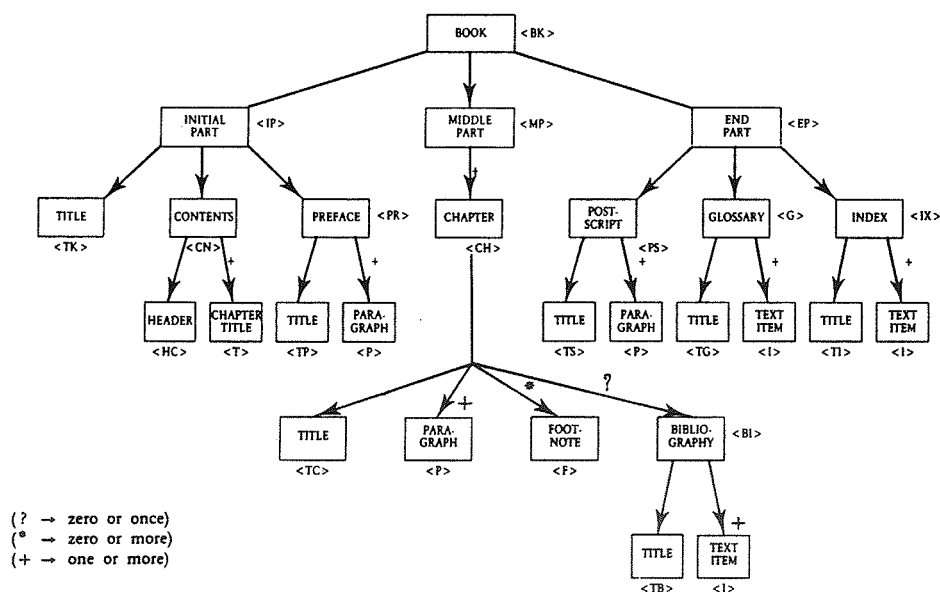
*Figur 1.* Dokumenttypedefinition (DTD) for en fagbog, jf. figur 2. (Efter FORMEX 1985)

---

DTD'en i figur 1 skal forstås således, at hvert element (venstre kolonne) består af ("genskrives som") de elementer, der står i parentes i højre kolonne, i den anførte rækkefølge. #PCDATA symboliserer den egentlige tekst. Efter et elementnavn på højre side kan der stå et plus (+), en stjerne (\*) eller et spørgsmålstegn (?); de har følgende betydning:

- + Elementet skal forekomme mindst én gang, men kan forekomme flere gange efter hinanden;
  - \* Elementet kan udelades, eller det kan forekomme én eller flere gange efter hinanden;
  - ? Elementet kan udelades, eller det kan forekomme højst én gang.
- Et umarkeret element skal forekomme netop én gang på den angivne plads.

De samme symboler er anvendt i den grafiske afbildning (figur 2) af (fag)bogens struktur. Fx består et kapitel <CH> af én titel <TC>, flere (+) afsnit <P>, måske nogle (\*) fodnoter <F>, og måske én (?) bibliografi <BI>.



Figur 2. Struktur af en fagbog, jf. figur 1. (Fra FORMEX 1985)

## SGML og ordbøger

Som gennemgående eksempel i det følgende anvendes artiklen *afgjort* (figur 3) fra en dansk-engelsk ordbog (Axelsen 1984), som jeg har konverteret fra fotosatsdata til SGML-

strukturerede data, som dernæst er importeret til GestorLEX redigeringsystemet med henblik på udarbejdelsen af en ny udgave.

---

**afgjort** *adj* (som er gået i orden) settled; (*udpræget*) definite, **F** decided (fx advantage, improvement); (*om person*) decided (fx she was very decided); *adv* definitely, **F** decidedly (fx better); unquestionably (fx he is unquestionably the best man); *en ~ sag* a settled thing; *det er (så godt som) ~ at* it is (as good as) settled that.

Figur 3. Artiklen *afgjort* fra Jens Axelsen: Dansk-engelsk Ordbog (Axelsen 1984)

---

Ordbogen flyttes altså fra sætteriets tekstbehandlingssystem, hvis typografiske kodning (figur 4) skal omsættes til en generisk. Samtidig har redaktionen udtrykt ønsker om, at præsentationen af ordbogens oplysninger bliver mere eksplicit.

---

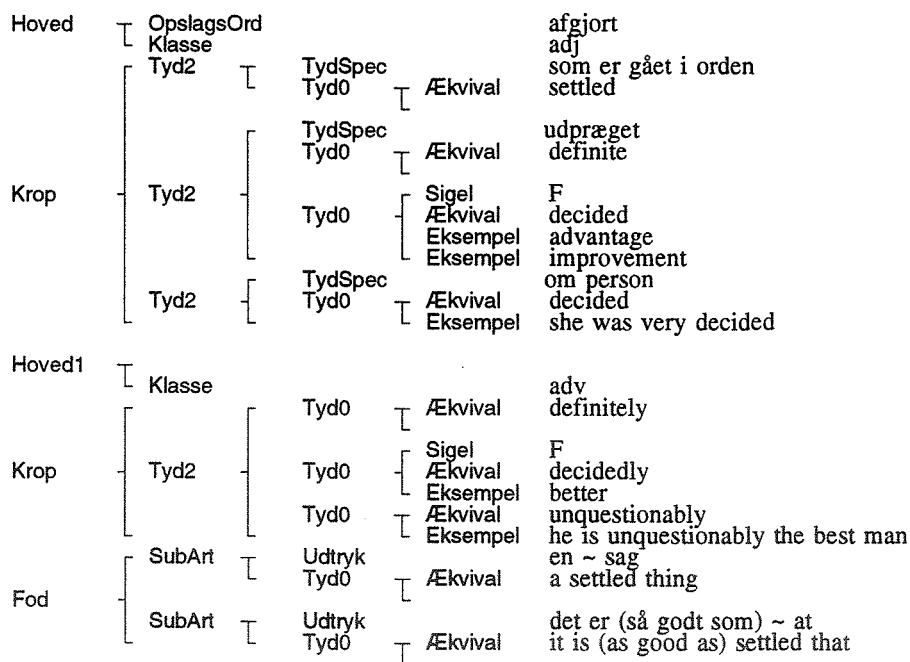
FED	<b>afgjort</b>		(fx
KURSIV	<i>adj</i>	ORDINÆR	she was very decided
	(		);
KURSIV	<i>som er gået i orden</i>	KURSIV	<i>adv</i>
	)	ORDINÆR	definitely
ORDINÆR	settled		,
	; (	GROTESK	<b>F</b>
KURSIV	<i>udpræget</i>	ORDINÆR	decidedly
	)		(fx
ORDINÆR	definite	ORDINÆR	better
	,		);
GROTESK	<b>F</b>	ORDINÆR	unquestionably
ORDINÆR	decided		(fx
	(fx	ORDINÆR	he is unquestionably the
ORDINÆR	advantage		best man
	,		);
ORDINÆR	improvement	KURSIV	<i>en ~ sag</i>
	); (	ORDINÆR	a settled thing
KURSIV	<i>om person</i>		;
	)	KURSIV	<i>det er (så godt som) ~ at</i>
ORDINÆR	decided	ORDINÆR	it is (as good as) settled that

Figur 4. Ordbogsartiklen opdelt i strukturmarkører og egentlig tekst.

---

Figur 5 viser en analyse (blandt mange mulige) af ordbogsartiklens opbygning. Grundlæggende består en ordbogsartikel af et **Hoved** med oplysninger om selve opslagsordet (her: ordklasse; men det kunne også være fx bøjning og udtale); en **Krop** med oversættelser af selve opslagsordet; og en **Fod** (andre ordbogsprojekter kalder dette afsnit for **Hale**) med mere eller mindre faste ordforbindelser, herunder egentlige idiomatiske udtryk. I det her viste eksempel er der to **Kroppe** og et indskudt "**Hoved1**". Det skyldes et redaktionelt ønske om, at ny ordklasse skal udløse ny artikel. På et senere stadium i konverteringsprocessen vil der fremkomme to adskilte artikler hver med sit homografnummer.

I den oprindelige artikel (figur 3) er de forskellige oversættelser af "afgjort" adskilt af komma eller semikolon; ofte er de desuden mærket med en **TydSpecifikation** (kursiv tekst i parentes) eller en **Sigel**, en versal i særlig skrift (grotesk). I eksemplet forekommer **F** = formelt sprog; i øvrigt bruger ordbogen bl.a. **S** = slang og **T** = talesprog. Der er grund til at antage, at variationerne (komma el. semikolon; TydSpec/Sigel el. ingenting) antyder et semantisk hierarki: de forskellige **Tyd** (en term, som skyldes Otto Jespersen) har betydningsmæssigt mere eller mindre med hinanden at gøre.



Figur 5. Strukturanalyse af artiklen *afgjort*

Efter aftale med ordbogens redaktør er det besluttet at strukturere tydene således: De enkelte **Tyd2** i en artikel har ikke meget til fælles; i den kommende ordbogsudgave skal flere **Tyd2** efter hinanden adskilles med semikolon og nummereres med arabertal. En **Tyd2** kan omfatte en eller flere **Tyd0**, som semantisk er nærmere beslægtede; er der flere **Tyd0** under samme **Tyd2** adskilles de blot med semikolon. Det flertydige komma afskaffes altså som separator

mellem forskellige tyd. Den opmærksomme læser vil savne en Tyd1. Den er foreløbig ikke taget i brug, men er reserveret til en eventuel underinddeling af Tyd2, typografisk fx markeret med a), b), c) ...

Processen som overfører de typografiske data til SGML-formaterede data indebærer bl.a. en klargøring af de typografiske data, jf. Norling-Christensen (1988), samt en parsing, som er analog med maskinel syntaks-analyse af naturligt sprog. Den beskrives nøjere i Norling-Christensen (1992); her skal blot antydes, at hovedreglen for skift til ny Tyd2 er, at der både forekommer et semikolon som separator og en TydSpecifikation eller en Sigel. Hvis der på TydSpecifikationens plads optræder en ny ordklasseangivelse, oprettes der derimod en helt ny artikel; for at fuldstændiggøre denne kopieres opslagsordet dertil. Hoved1 i figur 4 erstattes altså med et normalt Hoved bestående af opslagsordet "afgjort" og ordklassebetegnelsen "adv".

Filer Redigér Udclip Søg Præsentation Design Valg Hjælp	
Hom	casting; 5 (af ordre) placing;
Artikel	6 (kem.) liberation.
Hoved	
OpslagsOrd	afgjort
Klasse	adj.
Krop	
Tyd2	som er gået i orden
Tyd2	udpræget
Tyd2	om person
Artikel	
Hoved	
#NOTE	Ny homograf oprettet
OpslagsOrd	afgjort
Klasse	adv.
Krop	
Tyd2	definitely
Fod	
FodAfsnit	
SubArt	en ~ sag
SubArt	det er (så godt som) ~ at
Hom	afglans
	afglans (en) reflection (fx
DaEnRød Hom	
afgjort	

Figur 6. Skærmbillede fra redigeringsystemet GestorLEX. I strukturvinduet t.v. ses artiklerne *afgjort* delvis udfoldet (til og med niveau 4). Tekstvinduet t.h. viser ordbogsteksten i et typografisk format, som automatisk genereres ud fra strukturen og et sæt præsenteringsregler.



Processen omfatter også en konvertering af forkortelser etc. til en vedtagen standard; den indebærer bl.a. at alle forkortelser nu afsluttes af forkortelsespunktum. Resultatet af den automatiske proces ses i figur 6, som samtidig illustrerer, at det kun er forsvarligt at gennemføre automatiske ændringer af ordbøger, hvis der efterfølgende foretages en menneskelig kontrol. Bemærk nemlig, at **Foden** udelukkende rummer prøver på adjektivets brug; den skal altså flyttes op, hvor den hører til. I GestorLEX vil dette dog kun kræve nogle få tastetryk. For at sikre, at redaktøren er opmærksom på, hvor mekanikken har ændret hans tekst, sættes der noter ind, som han kan søge efter og tage stilling til.

### Ordbogsartiklen som SGML-struktur

Efter opdelingen i to artikler med samme opslagsord (homografer) kan artikelstrukturen beskrives med den DTD, som vises i figur 7. Det overordnede element **Hom** består af én eller flere **Artikel**; hvis der er flere, nummererer systemet dem med højststående arabertal (homografnumre). I øvrigt er elementnavnene de samme som på figur 5.

---

```

<!ELEMENT Hom (Artikel+) >
<!ELEMENT Artikel (Hoved, Krop, Fod?) >
<!ELEMENT Hoved (OpslagsOrd, Klasse) >
<!ELEMENT (Opslagsord, Klasse) (#PCDATA) >
<!ELEMENT Krop (Tyd2+) >
<!ELEMENT Tyd2 (TydSpec?, Tyd0) >
<!ELEMENT Tyd0 ( (Sigel | TydSpec)?, Ækvival, Eksempel*) >
<!ELEMENT (TydSpec, Sigel, Ækvival, Eksempel) (#PCDATA) >
<!ELEMENT Fod (SubArt+) >
<!ELEMENT SubArt (Udtryk, Tyd0+) >
<!ELEMENT Udtryk (#PCDATA) >

```

*Figur 7.* En (forenklet) dokumenttypedefinition (DTD) for Dansk-engelsk Ordbog. Den lodrette streg i "(Sigel | TydSpec)" betegner "enten/eller". Flere elementnavne på venstre side i en genskrivningsregel, fx "(Opslagsord, Klasse)", er blot en forkortet skrivemåde for, at disse elementer genskrives på samme måde.

---

SGML er som nævnt en international standard. Det indebærer, at ordbogsdata, som er lagret i SGML-format, ikke blot kan importeres til, benyttes i, og eksporteres fra GestorLEX-systemet. De kan også udveksles mellem forskellige ordbogsprojekter, og de er lette at bearbejde datamatisk, også uden for GestorLEX. En nøjere beskrivelse af GestorLEX falder iøvrigt uden for denne artikels rammer. Her skal blot nævnes, at systemet er detaljeret beskrevet i TEXTware (1991), og at det opfylder praktisk taget alle de krav som opstilles i DANLEX (1987:239-251).

**Litteratur**

- Axelsen, Jens. 1984. *Dansk-engelsk Ordbog*. 9. udgave (7. oplag 1991). København.
- Axelsen, Jens. 1990. *Dansk-engelsk Ordbog*. 1. elektroniske udgave. København.
- DANLEX. 1987: The DANLEX Group (Ebba Hjorth, Jane R. Jacobsen, Bodil Nistrup Madsen, Ole Norling-Christensen, Hanne Ruus): *Descriptive Tools for Electronic Processing of Dictionary Data. Studies in Computational Lexicography*. Lexicographica Series Maior 20. Tübingen.
- FORMEX. 1985: *Formalized Exchange of Electronic Publications. Standard generalized mark-up language (SGML) as described in Appendix B of the FORMEX manual*. Luxembourg, Office for Official Publications of the European Communities.
- ISO 8879. 1986. *International Standard ISO 8879. Information Processing - Text and Office Systems - Standard Generalized Markup Language (SGML)*. Genève.
- Norling-Christensen, Ole. 1988. Læsning af maskinlæsbare tekster. I: *Nordiske Datalingvist-dage og Symposium for datamatstøttet leksikografi og terminologi 1987. Proceedings*. Institut for Datalingvistik, Handelshøjskolen i København.
- Norling-Christensen, Ole. 1992. *Parsing Dictionary Data*. Et oplæg til Workshop on Tools and Methods for Practical Lexicography, Fifth EURALEX International Congress 4.-9. august, Tampere, Finland. (Under udarbejdelse).
- TEXTware. 1991. *LexWrite Brugermanual*. Version 2: April 24, 1991. København. Og *LexDesign Brugermanual*. Version 2: April 29, 1991. København.

Merete K. Jørgensen

## Ordbogsartikler som skriver sig selv - ? Stednavne som kilde i Gammeldansk Ordbog

I bestræbelserne på at beskrive det danske sprog i middelalderen så præcist og så fuldstændigt som muligt har der ved Gammeldansk Ordbog (GO) været gennemført et projekt med det formål at få afklaret, efter hvilke principper det ville være muligt at udnytte sproglige vidnesbyrd fra *stednavne* - en kildegruppe, som ikke tidligere har været systematisk udnyttet ved GO. Nærværende artikel beskriver, hvorledes ordbogsartikler automatisk kan genereres i GOs ordbogsredigeringsystem GestorLEX på baggrund af et maskinlæsbart stednavnemateriale, således at ord, om hvis eksistens kun stednavnebelæg vidner, også kan inddrages og behandles i GO.

Det følgende beskriver de erfaringer, som er opnået under forsøgene med automatisk dataudveksling mellem Gammeldansk Ordbogs redigeringsystem GestorLEX og en database indeholdende stednavnemateriale. De første arbejder med denne udveksling har behandlet de gammeldanske ord som kan ekstraheres fra stednavne, men som ikke findes belagt som gloser i andre gammeldanske kilder.

GO er en videnskabelig ordbog, hvis opgave det er at beskrive det danske sprog i perioden 1100-1515 så præcist og så fuldstændigt som muligt. For at opnå dette, udnytter GO i princippet hele den direkte overleverede tekstmasse fra middelalderen, overleveret enten i form af håndskrifter eller i form af trykte tekster. Ud over dette kernestof findes der imidlertid andre sprogkilder, som indirekte kan skaffe os viden om det middelalderlige sprog, og dermed kan supplere og nuancere den viden vi i forvejen har om ordforrådet i den gammeldanske periode. En af disse sprogkilder er stednavne.

I stednavne kan der findes dokumentation for ord, der ellers ikke er kendt fra middelalderlige kilder, og sproglige vidnesbyrd fra stednavne kan supplere den viden om et bestemt ords kronologi, lydlig udvikling, morfologi, geografiske distribution og betydningsindhold, der kan opnås gennem ordets ikke-proprie forekomst. For at få klarlagt under hvilke forudsætninger og på hvilken måde sprogstof ekstraheret fra stednavne kunne udnyttes i GO, blev der gennemført et projekt kaldet "Det gammeldanske ordforråd i stednavne - til brug ved Gammeldansk Ordbog". Der skal her redegøres for projektets overvejelser vedrørende anvendelse af ordstof fra stednavne som dokumentation for eksistens af ord, som ikke kendes fra periodens ikke-proprie ordforråd. Projektet samt dets konklusioner findes beskrevet i Jørgensen (1989).

Det må for disse navne først sandsynliggøres, at navnets dannelse ikke er sket efter 1515, som udgør GOs grænse. Dette krav kan honoreres fx ved middelalderlig overlevering. Når et navn opfylder dette krav og samtidig anses for at være dannet af et ord, som ikke ellers er kendt fra gammeldansk periode, så kan og skal dette ord inkorporeres i GO, under forudsætning af at ordet samtidig er kendt i ældre nydansk. Et sådant navnebelæg viser nemlig, at det pågældende ord på ét eller andet tidspunkt før 1515 har eksisteret i det danske sprog, og da det er bevidnet i det ikke-proprie kildemateriale i perioden efter 1515 må det

også have eksisteret i den gammeldanske periode. Det kan nemlig ikke godtages, at et ord eksisterer i sproget i en (før-gammeldansk) periode, derefter helt forsvinder i gammeldansk, for atter at genopstå eftermiddelalderligt. Den eneste acceptable antagelse er, at ordet har eksisteret i gammeldansk, men blot ikke er belagt i kilder fra perioden.

Et af projektets delformål var at afprøve metoder til (delvis) automatisk excerpering af stednavne, og som kildemateriale valgtes bind 12 fra serien *Danmarks Stednavne* (Bjerrum 1964). *Danmarks Stednavne*, som udgives af Institut for Navneforskning ved Københavns Universitet, behandler på videnskabeligt grundlag stednavne. De enkelte artikler har en fast og stringent struktur og de besidder tilstrækkeligt karakteristiske træk til at lade sig beskrive formelt; i projektet anvendtes DANLEX-gruppens taksonomi og formatbeskrivelse<sup>1</sup>.

**Skablund**, lb [D *sgå·blæn*, EJ *skå·blæn*, M *sgå·blīn*, *sgå·blūn*, *sgo·blīn*]. 23/6 1438 Skoblingh; Jb 1544 Skobling; KancBrevb c. 7—12/12 1551 Skoubylund; Jb c. 1571 Skublinge; KancBrevb 26/12 1580 Skubling; Jb 1582 Skoubølling, Skowbølling; Regnsk 1582 Skowbølling, Skobling; KancBrevb 12/1 1582 Skobling; KronSk 4/4 1583 Skoblinggaard; KancBrevb 14/5 1584 Skublinggaard; Jb 1589 Skoblingh, Skoblinggaardt; Mandt 1610 Sekobelingh; M 1664 Schoblinggaard; M 1688 Schabling; DATl, VSK Orig Skabling; M 1844 Skrablund; Da. Byremser p. 171 Skabling; Gst, Postadr 1958 Skablund. — Navnet er sandsynligvis et opr. \*Skōghbōling sammensat af subst. skov, gl̥da. skōgh, m. og det hyppigt forekommende efterled bōling 'beboer'; se iøvrigt p. 65 under Skåbling.

Figur 1

Fig. 1 viser, at den enkelte stednavneartikel kan indeholde følgende informationskategorier:

- opslagsnavn,
- angivelse af lokalitetens art,
- udtalekompleks,
- belægskompleks,
- tolkningsafsnit.

Typografisk er det især af vigtighed:

- at de enkelte kilders belæg med samt deres datering afgrænses af et semikolon
- at de(t) gammeldanske ord, som stednavnet anses at være dannet af i tolkningsafsnittet er markeret ved angivelse af *gl̥da.*, eller ved angivelse af \*, eller ved hjælp af spatium (eller en kombination af disse muligheder)<sup>2</sup>.

Det udvalgte bind blev gjort maskinlæsbart og lagret i en til projektet programmeret database. Indlæggelsen i denne base skete på baggrund af en nøje udarbejdet struktureret syntax-analyse over artiklernes format.

Med udgangspunkt i det nu strukturerede materiale som denne database udgjorde, kunne der foretages ordsøgninger med meget præcise søgekriterier, som bl.a. kunne indeholde krav til navnets overlevering og til det eller de ord, det ansås at være dannet af.

Opfølgende arbejde på projektets konklusioner var forsøg med direkte udveksling mellem projektets database og GOs editeringssystem GestorLEX. Til dette etableredes et samarbejde med TEXTware, det software-firma, som har udviklet GestorLEX.

På TEXTware parsedes en tekstfil, som bestod af en udlæsning af den gamle bases data. Denne parsing konverterede tekstens gamle præsentation til SGML-format<sup>3</sup>. Som det blev beskrevet ovenfor, besidder artikler fra Danmarks Stednavne en meget præcis struktur, der muliggør at opdele den i informationskategorier og lagre disse i en database. De samme fremtrædelsesformer blev nu benyttet til at mærke tekstfilen i SGML.

Formålet med denne første parsing var at få udskilt alle de artikler, som for det første indeholdt mindst et belæg før 1515, og som for det andet havde forekomst af tekststrengen *gl̄da.* og/eller \*. Det første kriterium sikrede, at resultatet ville blive alle de artikler, som behandlede navne, der var dannede senest i middelalderen, og dermed også sprogstof, som ikke var yngre end den gammeldanske periode. Formålet med det andet kriterium var at få mærket de ord som navnet ansås dannet af; af disse ord skulle de med \* mærkede senere danne opslagsord i ordbogen.

Det var imidlertid nødvendigt, før en kodning i SGML kunne foretages, nøje at overveje, på hvilken måde de enkelte informationskategorier fra stednavnematerialet skulle korrespondere med GOs grammatik.

Fig. 2 viser et udsnit af GOs grammatik. Som tidligere anført skulle de ord, hvoraf navnet ansås at være dannet, indgå som opslagsord, her benævnt *Opslag*; det som i GOs grammatik benævnes *Citat* skulle korrespondere med belægget fra stednavneartiklen; og *Kilde* til et sådant "citat" blev belæggets datering.

Følgende korrespondancer fastsattes således:

<b>GO</b>	<b>DS</b>
Opslag	tolket ordstof
Citat	belæg
Kilde	datering

Med udgangspunkt i en sådan konverteringstabel foretoges mærkningen af tekstfilen i SGML.

Fig. 3 viser et eksempel på en artikel kodet i henhold til SGML-standard. Under import i GOs grammatik genereres der et felt *Artikel*, hvortil der knyttes en note, som giver oplysning om det pågældende stednavns lokalisering. Der genereres et felt *Hoved*, hvortil der knyttes en note, som fortæller om ordet er valgt pga. forekomst af *gl̄da.* eller \*. Der genereres et felt *Opslag*, der indeholder det gammeldanske ord; i eksemplet her står 'grim', med langt i, men koderne for karakterer med diakritiske tegn vil senere blive konverteret efter GOs konventioner. Efterfølgende genereres et tomt felt *Ordkl*, hvis indhold senere skal indtastes. Endelig genereres en *Krop* og en *Krop 1* og et tomt felt *Tyd*, hvor betydningsbeskrivelsen senere kan indføres, og sluttelig genereres feltet *DokTyd*, som består af *Citat* (her: Grynskiær) og *Kilde* (her: 9/3 1311 (Århusb 1428-c. 1450 SRD VI 481)). Hele belægskomplekset (belæg + datering) lægges ind både i *Citat* og i *Kilde*, og det overflødige fjernes så senere.

## GOs grammatik

```

<!DOCTYPE Ordbog [
<!ELEMENT Ordbog                (Artikel+)>

<!ELEMENT Artikel                (Hoved, Hals*, Krop+, Hale?)>

<!ELEMENT Hoved                (HomNr?, Opslag, Udnyttelse?, Ordkl, Var*)>
<!ELEMENT HomNr                 (#PCDATA)>
<!ELEMENT Opslag                 (#PCDATA)>
<!ATTLIST Opslag                 antal NUTOKEN      #IMPLIED>
<!ELEMENT Udnyttelse            (#PCDATA)>
<!ELEMENT Ordkl                 (#PCDATA)>

<!ELEMENT Krop                  (Krop1 | Krop2)>

<!ELEMENT Krop1                (Tyd, Henv?, Stil?, Brug?, DokTyd+)>
<!ELEMENT Tyd                   (#PCDATA)>
<!ELEMENT Henv                  (JfOrd, Refer+)>
<!ELEMENT Stil                  (#PCDATA)>
<!ELEMENT Brug                  (#PCDATA)>
<!ELEMENT DokTyd                (Citat?, DokParal*, Kilde, DokForl?)>
<!ELEMENT Kilde                 (KildeK, KildeV+)>
<!ELEMENT KildeK                (#PCDATA)>
<!ELEMENT KildeV                (#PCDATA)>

```

Fig. 2. Udsnit af GOs grammatik; de understregede felter angiver de elementer som automatisk genereres.

```

</Artikel>
<#NOTE>ID=Grimskær    Naturnavn</> <Hoved><#NOTE>stjerne</>
<Opslag>gr<u>i</u>m</>
<Ordkl></></>
<Krop><Krop1><Tyd></>
<DokTyd><Citat>Grynskiær 9/3 1311 (Århusb 1428-c. 1450 SRD VI
481)</>
<Kilde><KildeK></>
<KildeV>Grynskiær 9/3 1311 (Århusb 1428-c. 1450 SRD VI
481)</DokTyd>

```

Fig. 3 Eksempel på SGML-kodet artikel.

Til konverteringstabellen skal således tilføjes:

GO	DS
[note t. art.]	lokalisering
[note t. Hoved]	* / gl.da.

Fig. 4. Ordbogsartiklen tegh, sb.; struktur- og tekstvindue.

LexDevelop - ok1sted (test)

Filer Redigér Udklip Søg Præsentation Design Valg Hjælp Afprøvning

Artikel

#NOTE ok

#NOTE ID=Tejbøl 2219

Hoved

#NOTE stjerne

Opslag te:gh

Ordkl sb.

Krop

Krop1

Tyd

#NOTE svarende til ænyda. tey, jf. Kalkar u. Teje og Skovteje; jf. ds 12 s. 171

*jordstykke*

DokTyd

Citat Teebel

Kilde

KildeK ds.

KildeV 8/10 1471 (LDipl)

DokTyd

Citat Teybøll

Kilde

te:gh sb.

*jordstykke* Teebel ds 8/10 1471 (LDipl), Teybøll ds 16/5 1486, Deybøll ds 3/2 1489 (c. 1550), theybøl ds 21/2 1492 (17. årh.), Tiebøll ds 18/5 1512.

(TH)o sb.

ler Thom ds 1338 (1558 ÆDA II 38), Thom sogen ds 1391 (1558 ÆDA II 37), 1438 (1558 ÆDA II 36), 1445 (1558 ÆDA II 41), 1446 (1558 ÆDA II 14), 1449 (1558 ÆDA II 37), Theem S. ds 6/4 1463, Thom ds 1467 (1558 ÆDA II 38), 8/6 1470, Thom S.

Én artikel fra Danmarks Stednavne kan godt afstedkomme oprettelse af flere artikler i GO ved import i GestorLEX, da det er antallet af ord mærket med *glda.* eller \* i tolkningsafsnittet, der bestemmer antallet af oprettede artikler, jf. fig. 1.

I det ordbogsdokument, som blev resultatet efter import af SGML-filen i GOs grammatik, skulle nu de artikler udsøges, som var blevet oprettet, fordi der forekom en asterisk i tolkningsafsnittet, altså fordi navnet af stednavneforskere ansås at være dannet af et ord, som ikke findes belagt i middelalderlige kilder. Artiklerne udvalgte ved hjælp af en søgeprofil, udformet i et SQL-lignende sprog<sup>4</sup>, som medtog alle artikler, som indeholdt den ovenfor omtalte note til feltet *Hoved* med 'stjerne' som data. Resultatet af denne søgning blev derefter nøje gennemgået med det formål at sikre, at de kunne honorere kravene til indlemmelse i GO. Mange kunne sorteres fra, fx fordi \* gjaldt en fællesgermansk, en gammelsvensk eller indoeuropæisk form eller fordi navnet mentes dannet af et andet sted- eller personnavn. Andre igen blev udeladt, da det ikke kunne godtgøres, at det pågældende ord fandtes eftermiddelalderligt. Som eksempel på en artikel i GO, som automatisk er genereret på baggrund af stednavnemateriale ses i fig. 4 artiklen *tegh*, sb.

Fig. 4 viser artiklen, som den fremstår på skærmen, i strukturvinduet til venstre, med artikelstrukturen synlig, og i tekstvinduet til højre, sådan som præsentationen tænkes at blive i det trykte medie. Ordklasse og betydningsbeskrivelse er blevet udfyldt, og kilde og citat adskilt. 'ds' foran kilden angiver belæggenes proveniens.

Antallet af nye artikler som blev oprettet under forsøgene med det enkelte bind af Danmarks Stednavne er hverken nedslående lille eller overvældende stort; og det skal sammenholdes med den forholdsvis lille mængde stednavne, som indgik i disse første forsøg. Afslutningsvis må det indrømmes, at ikke alle relevante oplysninger fra stednavnematerialet kom med. Grunden kan være mangelfulde oplysninger fra forlægget, som her i fig. 5, stednavneartiklen *Lømsgård*.

**Lømsgård**, gd [D *løms'go-r*, M *løms'go-i*, *lens'go-*]. Dag? 1474 (18. årh. efter Dyrskjøet Rep 3574) Lousgaard; 20/6 1485 (HdDb 1599) thoe gaarde kallis Lømsgaarde; dag? 1492 (18. årh. efter Dyrskjøet Rep 7333) Lousgaard; KancBrevb 23/6 1593 Lømsggaard; M 1664 Lømsggaard (en liden Hgd.); M 1688, DAtl Lømsggaard; VSK Orig Lynsggaard; Gst 1873 og 74, Trap<sup>4</sup>, Gst Lømsggaard; Postadr 1955 Lehnsggaard. — Forleddet er antagelig gen. af trænavnet løn, oldn. hlynr, m. 'ahorn'. Udtalen med [œ], [e] er i så fald vidnesbyrd om, at det østjyske område, der i nutiden har æ, e svarende til *glda.* y foran n f. eks. i subst. søn 'søn', ifl. KortFm p. 65 „fra Skanderborg til Skagen“, tidligere har strakt sig længere mod syd.

Fig. 5. Artiklen *Lømsgård*, Danmarks Stednavne, bd. 12: 168.

Det angives her, at navnet er dannet af trænavnet løn og derefter meddeles den oldnordiske form. Ordet 'løn' er ikke overleveret i *glda.* - ældste belæg i Kalkar er fra 1561 - men dette til trods mangler \*, og derfor medtages denne stednavneartikel og ordet 'løn' ikke ved den før omtalte søgning. Også andre mangelfulde oplysninger i forlægget kan få indflydelse på resultatet, men arbejde med især de nyeste udgivelser af Danmarks Stednavne vækker håb om større konsekvens i angivelser og i artikelstruktur. Et andet fortrin som disse nyere udgivelser besidder frem for de ældre, er at de, som et resultat af moderne bogproduktion, foreligger i maskinlæsbar stand.



Forsøgene indtil nu har omhandlet de gammeldanske ord som kun er overleveret til os via stednavne. Men stednavnematerialet rummer - som ovenfor omtalt - også oplysninger om i forvejen kendte gammeldanske ords kronologi, lydlig udvikling, morfologi, geografiske distribution og betydningsindhold. Kommende forsøg skal vise, hvorledes også sådanne oplysninger kan optages i GOs korpus. Materialet må udsøges på basis af andre søgekriterier, og det program, som mærker datafilen og siden kan udskrive den i SGML-format, må defineres derefter.

### Noter

1. Se DANLEXgruppen 1987 for yderligere oplysning om klassifikation af ordbogsdata og beskrivelse af datastrukturer.
2. En lignende analyse af et andet bind af Danmarks Stednavne er foretaget af Ebba Hjort (Hjort 1985).
3. SGML er det formelle sprog ved beskrivelse af ordbogens struktur (ISO 8879).
4. SQL står for Standard Query Language, se fx Date 1986.

### Litteratur

- Bjerrum, Anders. 1964. *Stednavne i Århus og Skanderborg Amter*. Danmarks Stednavne bd. 12. København.
- DANLEXgruppen. 1987. *Descriptive Tools for Electronic Processing of Dictionary Data*. Lexicographica, Series Maior 20. Tübingen.
- Date, C.J. 1986. *An Introduction to Database Systems I*. 4. Edition.
- Hjorth, Ebba. 1986. Leksikografisk onomastik og onomastisk leksikografi. I: *Mange bække små. Til John Kousgård Sørensen på tresårsdagen 6.12.1985*: 95-107. København.
- ISO 8879: *Information Processing. Text and Office Systems. Standard Generalized Markup Language (SGML)*. International Standard Organization, 1986.
- Jørgensen, Merete K. 1989. *Det gammeldanske ordforråd i stednavne - til brug for Gammeldansk Ordbog*. Rapport til Statens Humanistiske Forskningsråd. J.nr. 15-6120. Utrykt.
- Jørgensen, Merete K. 1990. Projektet "Det gammeldanske ordforråd i stednavne - til brug for Gammeldansk Ordbog". I: *Selskab for nordisk filologi. Årsberetning 1987-1989*: 106-114. København.



Del 6



Göran Karlsson

## Empiriska synpunkter på tvåspråkig lexikografi

I artikeln ges en redogörelse för tillkomsten av *Stora svensk-finska ordboken* (SSFO) som åren 1982-1987 utkom i tre delar på inalles drygt 3 000 sidor. Arbetet inleddes i början av 1960-talet vid Finska institutionen vid Åbo Akademi. Det skedde i Finska Litteratursällskapets regi och med finansiellt understöd av Finlands undervisningsministerium. I redogörelsen behandlas bl.a. vissa innovationer i angivandet av uppslagsordens uttal och betoning samt i utformningen av fraseologien. I senare hälften av föredraget ges en rapport om det arbete på en nedbantad version av SSFO som författaren f.n. är sysselsatt med. Den nya versionen är planerad att omfatta c:a 65-70 procent av innehållet i SSFO, och författaren redovisar de olika utrymmesbesparande åtgärder genom vilka omfånget minskas utan att den lexikaliska informationen i samma grad går förlorad.

Under åren 1952-1962 publicerades *Nykysuomen sanakirja*, i många avseenden en motsvarighet till Olof Östergrens *Nusvensk ordbok*. Den utkom i sex delar, och jag skrev en kort anmälan av varje del i *Finsk Tidskrift*. Då den sista delen hade kommit ut antydde jag i min anmälan att tiden nu var inne att försöka få till stånd en utförlig och tidsenlig svensk-finsk ordbok. Jag hade ingen personlig avsikt med min antydan, men professor Toivo Vuorela vid Finska Litteratursällskapet tog fasta på den och frågade om jag var villig att åta mig uppdraget. Jag tvekade mycket, jag hade nyss blivit utnämnd till min tjänst vid Åbo Akademi, men uppgiften föreföll också på något sätt lockande och jag beslöt mig för att försöka. Uppdraget formulerades så att jag skulle utarbeta en moderniserad ordbok av ungefär samma omfång som Knut Cannelins ordbok *Svenskt-finskt Lexikon* (7:e upplagan, 1953) och med den som utgångspunkt. Det visade sig i alla fall mycket snart, att en ytterligare omarbetning av ett lexikon med en stomme från seklets början inte hade varit ändamålsenlig. Endast en helt ny ordbok, utarbetad från grunden, kunde antas tillgodose de många och avancerade behov som hade framträtt med växande styrka. Den tilltänkta ramen skulle ha varit för snäv för att ge rum dels för den mängd nya ord som är självskrivna i ett modernt lexikon, dels för en tillräckligt fyllig och nyanserad fraseologi som många hade saknat hos Cannelin. Det kan dessutom nämnas, att uppslagsorden hos Cannelin till en del var mycket föråldrade. Ett exempel är ordet *sjösvamp* (en föregångare till *tvättsvamp*) som i Svenska Akademiens Ordboks arkiv har sitt sista belegg just från Cannelin och sitt nästista från 1852, alltså från hundra år tidigare.

På basen av mina erfarenheter kunde jag skissera upp vissa idealförutsättningar för utarbetandet av en tvåspråkig ordbok. Redaktionen borde till sin numerär vara sådan att den uppställda målsättningen kan nås inom rimlig tid. Redaktörerna bör självfallet ha en gedigen språklig utbildning som gärna fick omfatta också lexikografi, och de borde helst ha modersmålskompetens i båda språken. Åtminstone en redaktör borde ha det ena språket och en annan det andra språket som sitt egentliga modersmål och vara expert på det. Då skulle man inte behöva anlita utomstående expertis annat än i fråga om eventuell fackterminologi och kanske någon annan speciell fråga, och det skulle underlätta arbetet i hög grad. Fast

anställning, statlig finansiering och en inte alltför rigorös tidtabell vore en optimal ram. Arbetet på SSFO utfördes under förutsättningar som i de flesta avseenden inte nådde upp till detta ideal.

Då tanken på en omarbetning av Cannelins ordbok hade övergetts, var det naturligt att hämta uppslagsorden för ett nytt lexikon i första hand från *Illustrerad svensk ordbok* av Bertil Molde (3:e upplagan, 1964), som gav nödiga förklaringar och exempel på ordens betydelser i den tidens moderna svenska. Senare fick vi betydande tillskott från *Svenska Akademiens ordlista* (10:e upplagan, 1973). Det centrala finska materialet kom till största delen från *Nykysuomen sanakirja*, som då nyss hade blivit komplett, men i någon mån också från tidigare svensk-finska och finsk-svenska lexika. Åtskilliga fackordböcker och specialordlistor utnyttjades för kompletterande tillägg. Ett omfattande material insamlades dessutom genom studium av tidnings- och annan text, inkluderande bl. a. tvåspråkiga broschyrer om de mest olika slag av produkter och service. Fraseologin ägnades en särskild uppmärksamhet. Exempel för detta ändamål erhöles dels genom lån från andra ordböcker, speciellt från *Svensk handordbok*, utgiven av Ture Johannisson och K.G. Ljunggren, dels genom parallellstudium av svensk litteratur och dess finska översättning. Särskilt om prepositionerna fick jag ett rikhaltigt och mångsidigt material genom att ge finskstuderande i uppgift att som studieprestationer ur ett visst svenskt textpensum plocka ut exemplen på användningen av en viss preposition och motsvarande uttryck i textens finska översättning. Det resulterade i långa artiklar om de flesta prepositioner.

När man påbörjar ett ordboksarbete, gör man naturligtvis klart för sig syftet med den planerade ordboken: vilken eller vilka målgrupper har man i tankarna, och hur omfattande och detaljerad lexikalisk information avser man att ge de tilltänkta ordboksbrukarna. En tvåspråkig ordbok, t.ex. en svensk-finsk, borde helst, om tillräckliga resurser fanns, utarbetas i två versioner, en avsedd för svenskspråkiga och en annan för finskspråkiga användare. Det är mycket och mångt som skulle utformas på olika sätt i de två versionerna, om möjlighet till en differentiering fanns. Detta gäller inte bara rent tekniska saker som t.ex. sättet att ange uppslagsordens uttal och valet av språk för förkortningar och andra allmänna anvisningar, utan också rent innehållsmässiga frågor. Som exempel kan nämnas att man i ett för svenskspråkiga användare avsett lexikon bör ta med som uppslagsord en hel del sammansättningar med en i stort sett ordagrann finsk motsvarighet, som för finnar är mer eller mindre självklara, men som för personer med annat modersmål kan vålla osäkerhet framför allt i fråga om förledens kasus, men ibland också när det gäller valet mellan sammansatt ord och ordfogning; jfr *arbetschef* työpäällikkö, *arbetsledare* työnjohtaja, *arbetsträning* työhönvalmennus, *arbetsbesparande* työtä säästävä, *arbetsvägran* työstä kieltäytyminen.

Om möjligheter inte finns att utarbeta två olika versioner av en och samma tvåspråkiga ordbok, får man ta till lösningar som på bästa möjliga sätt tillgodoser de båda språkgruppernas intressen. Vi utgick till en början ifrån att den svensk-finska ordboken inte nämnvärt skulle användas annat än i Finland, men på ett senare stadium då intresset för finskan och finskstudier ökade i Skandinavien och i synnerhet i Sverige och sverigefinskan hade blivit en faktor att räkna med, vidgades arbetsperspektivet till att beakta också dessa omständigheter. I Finland räknade vi helt naturligt med att majoriteten av ordbokens användare skulle vara finskspråkiga, och nödig bruksinformation gavs därför i alla frågor på finska och i sådana avseenden som ur svensk synpunkt bedömdes väsentliga dessutom på svenska.

I en hel del lexikografitekniska frågor tog vi modell av *Illustrerad svensk ordbok*, i andra stannade vi för egna lösningar.

För angivandet av betoning och ljudlängd i uppslagsorden utformades ett system som utgår ifrån att huvudtrycket i regel är omarkerat om det, som fallet oftast är, ligger på ordets första stavelse, vilket det praktiskt taget alltid gör i finskan. Stavelsens vokal är då vanligen *k o r t* framför dubbelkonsonant eller konsonantförbindelse, t.ex. **panna**, **fjärran**, **öppen**, i övriga fall *l å n g*, t.ex. **men**, **hagel**, **yla**, **fiende**, **snö**. I fall som avviker från denna grundregel markeras huvudtrycket med en punkt i radens övre kant. Punkten placeras direkt efter stavelsens vokal om vokalen är *l å n g*, t.ex. **a•ln**, **analo•g**, annars efter den bokstav eller bokstavsförbindelse som är tecken för närmast följande konsonantljud (och detta är då långt), t.ex. **men•** [menn], **artis•t**. Det nya i det här systemet är att en fristående accent är utsatt i själva uppslagsorden, också i enstaviga, och att en särskild uttalsparentes på så sätt oftast kan undvaras. Systemet är delvis i bruk också i *Svensk ordbok* (1986), i stor utsträckning i flerstaviga, men vid behov också i enstaviga ord, t.ex. *kapita'l*, *röveri'*, *a'ln*.

Markeringen av uppslagsordens uttal i övrigt skedde med användning av gängse finsk ortografi. Det innebär t.ex. att *sje*-ljudet har sitt finska tecken *š*, att bokstaven *u* är tecken för vokalljudet *o* och att lång vokal anges med dubbel bokstav, t.ex. **natio•n** [natšuu]. De två uttalen av bokstaven *o* i svenskan kommer till uttryck så, att *o* som sådant har fått bli tecken för *å*-ljudet som i finskan, t.ex. **ekono•m**, medan uttalet *o* markeras med en punkt under bokstaven, t.ex. **kø**. Genom att sätta ut en fristående betonings- och kvantitetsangivande punkt i uppslagsorden och genom att låta *o* vara tecken för *å*-ljudet har största delen av uttalsparenteserna i SAOL kunnat undvaras, inte bara vid talrika flerstaviga ord, utan också vid många enstaviga, t.ex. **flop•**; i SAOL: **flop** [flåpp].

Ytterligare följande specialtecken behövdes i uttalsparenteserna: *ε* för att ange det halvöppna *e*-ljudet t.ex. i **enque• te** [angkeet], *ʌ* för att beteckna svenskans *u*-ljud t.ex. i **rwan•dier** [r-ʌ-an-] och *œ* för att ange det *ö*-ljud som framför *r* är mycket öppnare i svenskan än i finskan, t.ex. **arrangö•r** [arangšœer].

Hur de enskilda ordartiklarna är uppbyggda skall inte här närmare beskrivas, det finns några prov på dem längre fram.

Ett troligen mindre vanligt förfarande att utgående från målspråket komplettera fraseologin förtjänar att kort beröras. Finskan är mycket rik på avledningar, det finns ett stort antal både nominala och verbala avledningsändelser. En av dem ingår t.ex. i *juoksija* 'löpare' och *kääntäjä* 'översättare'. Men många av dessa avledningar är alltför tillfälliga för att noteras som uppslagsord i lexika, och en stor del av dem har ingen motsvarighet i svenskan. En sådan är *osaaja* av *osata* 'kunna', som låter helt naturlig i ett uttryck som *ranskan osaajia tarvitaan lisää*, men som inte kan återges med ett obefintligt \**kunnare* och därför inte skulle återfinnas vare sig i en svensk-finsk eller en finsk-svensk ordbok, om man inte under **kunna** tog upp en fras av typen *det behövs flera (personer) som kan franska*. En annan central avledningstyp är de s.k. egenskapsnamnen på *-(u)us*, *-(y)ys*, oftast avledda av adjektiv som bl.a. *suoraselkäisyys* av *suoraselkäinen* 'rakryggad'. I svenskan talar man inte om \**rakryggighet*, och därför måste *suoraselkäisyys* närmast återges med *rakryggad hållning*, ett exempel som kan tas upp under någotdera av dessa ord med *suoraselkäisyys* som en bland andra motsvarigheter.

Grammatik i lexika är en fråga som kan diskuteras. I SSFO har hänsyn till målspråkets grammatik tagits bl. a. såtillvida, att objektssatser inledda av konjunktionen *att* tagits med i en utsträckning som ger möjlighet att mångsidigt belysa bruket av den konstruktion i finskan som kallas referativ satsförkortning. På samma sätt har ett representativt antal temporala *då*-satser tagits med under konjunktionen **då** för att illustrera olika sätt att använda finskans temporale satsförkortning.

Alla ordböcker är som vi vet mer eller mindre föråldrade redan när de kommer ut. Året innan del 3 av *Stora svensk-finska ordboken* gavs ut, dvs 1986, utkom *Svensk ordbok*, 11:e upplagan av *Svenska Akademiens ordlista* och *Nyord i svenskan från 40-tal till 80-tal*, och man kunde konstatera att en ny generation av svenska ordböcker hade sett dagen. Ett svensk-finskt lexikon baserat i huvudsak på tidigare svenska ordböcker, i och för sig förtjänstfulla, visade sig plötsligt vara mera efter sin tid än man hade önskat att det skulle vara. Tanken på en moderniserad och samtidigt komprimerad version i en volym aktualiserades snart, och det är arbetet på den jag nu är sysselsatt med.

Arbetet går till så, att en förstora kopierad kopia med breda marginaler tas av varje sida i SSFO. I marginalerna införs nödiga ändringar och tillägg, medan direkta strykningar görs med rödpenna i själva texten. Först noteras med särskilda tecken i kanten av varje spalt vilka uppslagsord i *Svensk ordbok*, SAOL 11 resp. *Stora svensk-engelska ordboken* (1988) som finns på kopian, och de som saknas understreckas med rött i resp. bok. De ord på kopian som inte är upptagna i någon av de tre böckerna blir nästan alla strukna, medan de övriga till största delen bibehålls, i synnerhet de som ingår i *Svensk ordbok* och SAOL eller i någondera av dem. Vid bearbetningen tillfogas större delen av de ord som blivit understreckade med rött i källböckerna och dessutom det mesta av det material som ingår i olika ordlistor producerade av Sverigefinska språknämnden. Värdefulla tillskott ger dessutom bl. a. informationsbladen *Språkbruk* och *Kieliviesti* som utges av Svenska språkbyrån i Helsingfors resp. Sverigefinska språknämnden.

För finskans del är en ny ordboksgeneration på kommande. I slutet av år 1990 utgavs första delen av *Suomen kielen perussanakirja*, en efterträdare till *Nykysuomen sanakirja*, som redan till stor del har blivit föråldrad. Två delar till av den nya ordboken är planerade att utkomma före utgången av 1994, det totala omfånget är beräknat till 2 000 sidor och antalet uppslagsord till inemot 100 000, varav över 20 000 sådana som inte ingår i *Nykysuomen sanakirja*.

Nya gedigna svenska och finska ordböcker skapar nu mycket goda förutsättningar för utarbetandet av en svensk-finsk och en finsk-svensk ordbok med modernt innehåll. Men språken utvecklas hela tiden. För finskans del har det beräknats att tre nya ord kommer till per dag, och någonting liknande gäller väl också svenskan. Det innebär drygt tusen nya ord i året, och det gäller att vaksamt följa med språkens utveckling och fånga upp så många som möjligt också av nykomlingarna.

Att banta ner innehållet i en ordbok går bara till en liten del rent mekaniskt; det nämndes redan att sådana uppslagsord som inte återfinns i någon av de tre källexika nästan automatiskt stryks. De är antingen föråldrade eller också ord från olika fackområden som kom att tas med i en utsträckning som ur nuspråkssynpunkt är omotiverat stor, då fackordböcker tillkommer på allt flera specialområden och förnyas i takt med utvecklingen. I huvudsak gäller det att finna på andra nedskärningsobjekt. Och detta har också visat sig vara möjligt utan att väsentlig lexikalisk information i alltför hög grad behöver gå förlorad.

Idealet att låta varje uppslagsord bilda ett eget stycke går inte att förverkliga i en utförlig tvåspråkig ordbok som har till syfte att med nödiga synonymer och en mångsidigt exemplifierande fraseologi belysa uppslagsordens användning, särskilt inte om lexikonet skall omfatta endast en volym. Avvikelserna från detta ideal är stora i SSFO och blir ännu större i dess förkortade version (här kallad SSFO 2). Bl. a. kommer uppslagsord som i SSFO bildar var sitt stycke ofta att föras ihop till ett. Som exempel kan nämnas att de två homograferna *by* sammanförs i ett stycke och att ett antal sammansättningar ytterligare tillfogas:



## SSFO

1. **by** -n -ar kylä; *Liela* ~ Lielahden kylä; *en avsides belägen* ~ syrjäinen t. syrjäkylä; *hela* ~n var i rörelse koko kylä t. kylän väki<sup>5\*</sup> oli liikkeellä; *också från andra* ~ar hade man kommit myös toiskyläläisiä oli tullut; *ute i ~n talas det mycket om henne* kylällä hänestä puhutaan paljon; *sätta* 1. *ställa kyrkan mitt i ~n* panna kirkko\* keskelle kylää; ISL: *skicka barn i 1. till ~ 1. till ~s(s)* och gå själv efter pane lapsi<sup>13</sup> asialle, mene itse perässä.
2. **by** -n -ar: (storm~) tuulenpuuska, puuska, vihuri<sup>2</sup>, tuulispää; *vindstyrkan i ~arna var hela 170 km/t* tuulen voimakkuus<sup>2</sup> oli puuskittain peräti 170 km/t.
- by•allmänn'ing** (kylän) yhteis- t. yleismaa, kylä(n)maa.
- byallag** ~et kyläkunta\*; (by) kylä; (kvarterförening) asukasyhdistys.
- byalman** ~nen kylänmies<sup>7</sup>; kyläläinen. -mål kyläosuus<sup>2</sup>. -rå s. kylänraja<sup>2</sup>.
- byllbebyggelse** kyläasutus. -bø ~n kyläläinen, kylänmies<sup>7</sup>, *det är ngn ~ som har gjort det här* tämän on tehnyt jku omankyläläinen t. oman kylän asukas<sup>3\*</sup>.

## SSFO 2

- by•** -n -ar 1 (storm~) tuulenpuuska, puuska, vihuri<sup>2</sup>, tuulispää; *vindstyrkan i ~arna var hela 170 km/t* tuulen voimakkuus<sup>2</sup> oli puuskittain peräti 170 km/t. 2 kylä; *Liela* ~ Lielahden kylä; *en avsides belägen* ~ ks avsides; *hela* ~n var i rörelse koko kylä t. kylän väki<sup>5\*</sup> oli liikkeellä; *också från andra* ~ar kom man myös toiskyläläisiä tuli; *så säger man ute i ~n* niin kylällä puhutaan; *ställa kyrkan mitt i ~n* ks kyrka; *skicka barn till ~(s) och gå själv efter pane lapsi<sup>13</sup> asialle, mene itse perässä.*
- allag** kyläkunta\*;  
(kvarterförening) asukasyhdistys.
- allmänn'ing** (kylän) yhteis- t yleismaa, kylä(n)maa. -almål lhist.l kyläosuus<sup>2</sup>. -bø kyläläinen, (byaman) lm.l kylänmies<sup>7</sup>; *det är någon ~ som har gjort det här* tämän on tehnyt joku omankyläläinen t. oman kylän asukas<sup>3\*</sup>.

Ordningsföljden mellan de två homograferna är motiverad av att sammansättningarna hör ihop med **by** 2. Sex olika stycken i SSFO blir i dess förkortade version kombinerade till ett. Genom att notera uppslagsord och ändelser endast *en* gång, genom att utelämna de mindre väsentliga sammansättningarna **bybebyggelse** och **byarå**, genom att något förkorta exemplet *ute i ~n talas det mycket om henne* och i samband med exemplen *en avsides belägen* ~ och *sätta kyrkan mitt i ~n* hänvisa till andra uppslagsord samt genom vissa andra smärre justeringar kan 22 lexikonrader i SSFO komprimeras till 14 i SSFO 2 (med något mindre tryckstil), utan att någon viktigare lexikalisk information går förlorad.

I tvåspråkiga ordböcker är det vanligt att inleda en ordartikel så att man efter uppslagsordet räknar upp motsvarigheterna i målspråket och vid behov preliminärt preciserar deras betydelser med tillhjälp av parentetiskt angivna synonymer till uppslagsordet och först därefter vid behov närmare belyser användningen av målspråkets synonymer med eventuellt erforderliga exempel. Avvikande från denna metod praktiseras i SSFO 2 i stor utsträckning ett nytt förfarande: direkt efter eventuella synonyparenteser följer exempel som i mån av

möjlighet är så formulerade, att uppslagsordets motsvarighet(er) uppträder i sin grundform; synonymparenteserna är ofta utspridda i ordartikeln så att de står just framför eller i samband med det eller de exempel, med vilket eller vilka resp. betydelse skall belysas. Ett par prov på detta förfarande jämfört med det gängse:

## SSFO

**erllägga** [kirjak.] maksaa<sup>5</sup>, suorittaa\*;  
~ böter, hyra, skatt maksaa<sup>5</sup> sakkoja, vuokraa, veroa; ~ en avgift suorittaa\* maksu; ~ likvid maksaa<sup>5</sup>; arbetsgivarens skyldighet att ~ försäkringspremier työnantajan vakuutusmaksuvelvollisuus<sup>2</sup>. -nde ~t maksaminen, suorittaminen, maksu, suoritus; mot ~ av en avgift på 10 mk 10 mk:n maksua vastaan, suorittamalla 10 mk:n maksu.

**erlnå** [kirjak.] saavuttaa\*;  
(åstadkomma) saada<sup>4</sup> aikaan; ~ goda resultat saavuttaa\* hyviä tuloksia; det enda man ~dde var att - - ainoa tulos oli että - -, ainoa mikä saatiin aikaan oli se, että - -; för ~ende av större likhet suuremman yhtäläisyyden saavuttamiseksi.

## SSFO 2

**erllägga** [el. äär-] ~ böter, hyra skatt maksaa<sup>5</sup> sakkoja, vuokraa, veroa; ~ en avgift suorittaa\* maksu; ~ likvid maksaa<sup>5</sup>; skyldighet att ~ försäkringspremie vakuutusmaksuvelvollisuus<sup>2</sup>; mot ~nde av en avgift på 50 mk 50 mk:n maksua vastaan, suorittamalla 50 mk:n maksu. -nå (åstadkomma): ~ goda resultat saavuttaa\* hyviä tuloksia; detta var det enda man ~dde tämä oli ainoa mikä saatiin aikaan t ainoa aikaansaannos t tulos; för ~ende av större likhet suuremman yhtäläisyyden saavuttamiseksi.

Systemet har förstås sina begränsningar. Det går bra att med **flintla** kombinera **-dolk**, **-glas**, **-gods**, **-hård**, **-kniv**, **-lås**, **-läsigevär** och **-redskap**, men att sedan ännu fortsätta med **-skalle** vore nog väl hårt!

Uppslagsord med gemensam förled kan ibland radas upp efter varandra utan översättning, varvid även motsvarigheterna i målspråket kan anges med gemensam förled, t.ex. **flyktingllhjälp**, **-invandring**, **-kommissarie**, **-kvot**, **-läger**, **-problem**, **-skara** pakolaislapu\*, -maahanmuutto\*, -komissaari, -kiintiö, -leiri, -ongelma, -joukko\*.

Strävan var i SSFO att för uppslagsorden så fullständigt som möjligt redovisa de finska motsvarigheter som i olika sammanhang kunde anses adekvata eller m a o i olika exempel kunde komma i fråga. Också för SSFO 2 gäller samma strävan. Men en del exempel har blivit inaktuella och strukits. Å andra sidan har ordens betydelseutveckling krävt ett antal nya exempel. I många fall har exemplen kunnat förkortas, dels genom strykning av element som visat sig mer eller mindre ovidkommande, dels genom att byta ut i sammanhanget irrelevanta satsdelar mot andra som tar mindre utrymme. Enkla tempus av verben kommer därför att vara vanligare än sammansatta, och A blir ett oftare återkommande subjekt än förut. Jag och min, du och din som ofta inte behöver översättas tränger kanske i någon mån ut han, hans och hon, hennes, och de singulara verbformerna får möjligen ett litet övertag över de plurala, vilkas finska motsvarigheter är något längre än de singulara.

Då ordboken skall omfatta endast *en* volym kan hänvisningar användas i större utsträckning än tidigare. Särskilt ordspråk och stående talesätt av olika slag behöver därför översättas bara under ett uppslagsord, och under de övriga hänvisas till detta. T.ex. för ordspråket *lyckan står den djärve bi* ges de tänkbara finska motsvarigheterna under ordet **djäv**, och under de övriga kommer en hänvisning.

Den strikt alfabetiska ordningsföljden för uppslagsorden frångås i en del marginella fall i syfte att spara utrymme. I samband med verbet **bröa** nämns sålunda även synonymen **bröda**, fastän den borde komma efter **bröd**; en rad fås på så sätt till godo utan att någon praktisk nackdel uppstår.

Systemet för markering av betoningen kompletteras i ett avseende. Huvudregeln, enligt vilken betoningen av lederna i en sammansättning är densamma som i de enskilda lederna, har ett undantag. T.ex. i det stycke som börjar med **all•tllintill** följer senare **-omfatt'ande** och **-uppslu'kande**, och i dessa och liknande fall är betoningen i senare leden en annan än då ordet uppträder som enkelt. Bitonen markeras här med en accent, och även förleden behöver då sin betoningsspunkt.

Ordklass och böjning hos nedannämnda grupper av uppslagsord anges med en förenklad beteckning på de sätt som följande exempel visar.

**arbetare** *s* i st. f. **arbetare** *-(e)n* ~ hos talrika substantiv på **-are** med samma böjning som *arbetare*.

**flitig** *a* i st. f. **flitig** *-t -are* hos den stora grupp adjektiv som har ändelserna *-t* och *-are* fogade till en oförändrad grundform.

**stanna** *v* i st. f. **stanna** *-de -t* hos den stora grupp verb som har ändelserna *-de* och *-t* fogade till en oförändrad infinitiv.

**blandning** *s* i st. f. **blandning** *-en -ar*

**besinn•ing** *s* i st. f. **besinn•ing** *-en*

**frihet** *s* i st. f. **frihet** *-en -er*

**behö•righet** *s* i st. f. **behö•righet** *-en*

Alla avledningar på **-ning**, **-ing** och **-het** har ändelsen *-en* i bestämd form i singularis och ändelsen i obestämd form i pluralis - om detta numerus är i bruk - är alltid *-ar* efter **-(n)ing** och *-er* efter **-het**.

Vilka av dessa avledningar som används i pluralis blir inte angivet i ordboken; ett eventuellt pluralbruk är inte heller alltid lätt att konstatera, och det torde kunna anses mindre väsentligt att notera i ett svensk-finskt lexikon än i ett finsk-svenskt. Många av dessa avledningar utelämnas, om deras finska motsvarigheter är få och lätta att bilda med tillhjälp av en översikt över deras bildningssätt som skall ingå i början av ordboken. För flertalet av de övriga begränsas antalet finska motsvarigheter på ett sätt som i varje enskilt fall visar sig lämpligt, och ordboksbrukaren får även då efter givna mönster själv bilda sådana som därutöver kommer i fråga. En fullständig uppräknings har i vissa fall visat sig bli oformligt lång.

En uttalsklammer kan inte sällan förkortas, t.ex. **flagran•t** [el. *-angt*] i st. f. [-ant el. *-angt*]. Ibland kan den undvaras, och om flera ord med samma uttal på lämpligt sätt förenas i ett och samma stycke, behöver uttalet markeras endast en gång, t.ex. **e•länllda** el. **elän•llda** *tpm. a ---. -de ~t ~n ---. -dig a ---.*

Av förkortningarna kan vissa utgå och andra ersättas med kortare, t.ex. *neutr.* > *n*, *taipum.* > *tpm.*, *tietok.* > *atk.* Förkortningen fraseol. *ks.* ersätts med →. Efter de förkortade

namnen på ordklasserna bortlämnas punkten, t.ex. *s, a, v, pron* osv, likaså efter *ks, t* och *vrt* [= katso 'se', tai 'eller', vertaa 'jfr'] som förekommer mycket ofta. Också i övrigt avlägsnas en hel del punkter och kolon.

Rätt mycket utrymme sparas genom att spärrad stil ersätts med kursiverad vid hänvisningar.

Erforderlig information betingad av de redovisade åtgärderna för nedbantning och utrymmesbesparing skall ingå i början av den nya ordboken. Sammanfattningsvis kan sägas att den, när den föreligger färdig, volymmässigt motsvarar 65-70 procent av SSFO, men tack vare en sträng ekonomisering med utrymmet procentuellt innehåller betydligt mera lexikalisk information än så. Arbetet har ännu (i maj 1991) inte avancerat längre än ett stycke in på bokstaven *f*, men takten ökar förhoppningsvis med rutinen, och om allt går väl kan ett slutresultat föreligga kring medlet av 1990-talet.

Hans de Groot

## Från svensk-engelsk till svensk-nederländsk

I september 1990 startades ett svensk-nederländskt/nederländsk-svenskt ordboksprojekt på Skandinaviska Institutionen vid universitetet i Groningen 1990. Till grund för den svensk-nederländska delen ligger de svenska ord och fraser som finns upptagna i *Stora svensk-engelska ordboken* (Esselte, 1988). Redaktionens erfarenheter av att omarbete en svensk-engelsk ordbok till en svensk-nederländsk är ämnet för denna presentation.

### En presentation av det svensk-nederländska och nederländsk-svenska ordboksprojektet

I september 1990 började arbetet på en svensk-nederländsk och en nederländsk-svensk ordbok på Skandinaviska Institutionen vid universitetet i Groningen. Det viktigaste skälet att starta projektet var att det inte fanns någon större nederländsk-svensk/svensk-nederländsk ordbok. Den senaste större ordboken utkom på 1950-talet (Boer-Den Hoed, 1956). Den är emellertid slutsåld sedan många år och dessutom föråldrad och därför inte längre användbar. Två fickordböcker finns att tillgå i handeln (Elsevier, 1984; Wolters, 1988), men dessa har ett litet omfång och ger bara ringa eller ingen information om ordens användning eller om uttal och böjning.

I dag måste de som seriöst sysslar med nederländska och svenska översätta via ett tredje språk, till exempel engelska. Det är omständligt, dyrt (man måste skaffa sig ett stort antal ordböcker) och det leder lätt till översättningsfel.

### 1. Pilotprojektet

Det var mot denna bakgrund som ett så kallat pilotprojekt startades på Skandinaviska Institutionen i Groningen 1989. Avsikten med det var att kartlägga möjligheterna att förverkliga en nederländsk-svensk/svensk-nederländsk ordbok. Pilotprojektet finansierades gemensamt av utbildningsdepartementen i Nederländerna och Sverige. Resultatet blev en rapport som skickades till de båda utbildningsdepartementen. I Sverige behandlades rapporten också av riksdagens utbildningsutskott (UbU18, s. 23-24).

I pilotstudien kartlades marknaden för en eventuell ordbok. Utgångspunkten var tanken att ordboken kunde säljas i fyra länder, nämligen Nederländerna (15 miljoner invånare), Belgien (6 miljoner nederländskspråkiga invånare i den flamländska delen av landet), Sverige (8,5 miljoner invånare) och Finland (0,3 miljoner invånare som har svenska som modersmål och ett obekant antal som behärskar svenska). Sammanlagt har dessa länder alltså minst 30 miljoner invånare som använder något av de två språken. Det största intresset för en ordbok finns troligen bland språkstudier vid de universitet som erbjuder undervisning i svenska resp. nederländska. Dessutom finns det behov inom folkbildningen, forskningen och

näringslivet, bland turister och översättare samt slutligen hos biblioteken, varav det finns ungefär ett tusen i såväl Nederländerna som Sverige.<sup>1</sup>

Den tänkbara upplagens storlek bedömdes i pilotstudien till ungefär 3-4000 exemplar, utgående från att ordboken skulle publiceras både i Nederländerna/Belgien och Sverige/Finland och att den första upplagan skulle säljas inom tre år. Den återkommande efterfrågan, dvs. efterfrågan som kommer tillbaka varje år, skulle förmodligen ligga på några hundratal exemplar.

I pilotstudien redogjordes också utförligt för kontakter som tagits med olika förlag i Nederländerna och Sverige. Rapporten avslutades med en kalkyl över kostnaderna, som efter de preliminära beräkningarna skulle uppgå till minst 400.000 nederländska gulden (SEK 1.200.000,-). Under våren 1990 kom beskedet att Sverige och Nederländerna var beredda att vardera ge ett finansiellt stöd till projektet på 200.000 gulden, så att arbetet på ordboken kunde sättas igång.

## 2. Ordbokens profil

Ordbokens viktigaste målgrupp utgörs av de studerande i det nederländsk- och svenskspråkiga området som seriöst sysslar med det nederländska/svenska språket och som inte eller ännu inte helt kan klara sig med enspråkiga ordböcker. Men även de som behärskar språket bättre ska ha nytta av ordboken, samtidigt som (å andra sidan) även brukare som inte är vana vid att arbeta med ordböcker ska ha användning för ordboken.

En kort beskrivning av ordboken i fem punkter:

- a. den utges i Nederländerna av Van Dale Lexicografie i Utrecht<sup>2</sup> och i Sverige av Norstedts Förlag; troligen får ordboken ett annat band i Sverige, men till innehållet är de nederländska och svenska versionerna lika;
- b. den ska bestå av två delar, en nederländsk-svensk och en svensk-nederländsk;
- c. den ska vara medelstor, dvs. varje del ska omfatta ca. 750 sidor och 45.000 uppslagsord per del;
- d. den ska ge inte bara ordens betydelse, utan också grammatiska uppgifter och information om uttal och användning;
- e. den ska behandla nutidsnederländskan och nutidssvenskan (en modern ordbok).

## 3. Uppläggning och arbetssätt

Projektet kan indelas i två faser. I den första produceras den svensk-nederländska delen. Utgångspunkten är det svenska material som finns i *Stora svensk-engelska ordboken* (Esselte, 1988). Alla uppslagsord, grammatiska uppgifter (ordklassstillhörighet), etiketter för stil eller ämnesområde och fraser i denna ordbok har matats in i redigeringsprogrammet EDI<sup>3</sup>. Det som redaktörerna gör är att bedöma om de vill ta upp de uppslagsord och fraser som finns i databasen och, om så är fallet, översätta materialet till nederländska. Ibland lägger de till något, fast inte ofta eftersom ordlistan är ganska uttömmande.

Under andra fasen av projektet, när den nederländsk-svenska delen kommer till stånd, är utgångspunkten en lista över nederländska ord och uttryck som ligger till grund för de nederländsk-främmande språk delarna i Van Dales handordbokserie.

Den nederländska listan innehåller ca. 45-50.000 uppslagsord. Om man använder samma definition på ett uppslagsord uppgår antalet uppslagsord i den svenska listan till ca. 75.000. Eftersom vår ordboks båda delar ska vara ungefär lika omfattande måste redaktörerna minska omfånget av det svenska materialet med minst 25 procent.

Det finns olika sätt att åstadkomma en sådan minskning av omfånget. Först och främst kan man stryka uppslagsord. Vissa kategorier uppslagsord kommer då särskilt i betraktande. Det finns t.ex. inga stora skillnader vad gäller sättet på vilket svenska och nederländska bildar sammansättningar. Därför behöver man inte ta upp så många sammansättningar som den svensk-engelska ordboken gör, troligen på grund av att de engelska sammansättningarna ofta inte är genomskinliga för svensktalande.

Detsamma gäller adverb. Det är meningslöst att t.ex. ta upp både adjektivet *hjaltemodig* (SvEng *heroical*, SvNl *heldhaftig*) och adverbet *hjaltemodigt* (SvEng *heroically*, SvNl *heldhaftig*).

Förutom genom att stryka uppslagsord finns det andra sätt att skära ned omfånget. I den svensk-engelska ordboken tas t.ex. ofta upp ett stort antal översättningar per uppslagsord. Det svenska ordet *indelning* får t.ex. fem översättningar (*division*, *subdivision*, *classification*, *grouping* och *mapping out*), medan vi tycker att det i den svensk-nederländska ordboken räcker med bara en översättning (*indeling*) som täcker det svenska ordet helt. I sig är det möjligt att lägga till ett par variantöversättningar (t.ex. *classificatie*). Redaktionens politik är emellertid att vara sparsam med dem, för det första eftersom vi tycker att svenskan och nederländskan är så pass lika att det inte behövs många varianter, för det andra eftersom variantöversättningar tar väldigt mycket plats i ordboken på grund av att de ska kommenteras så att användarna förstår i vilket avseende de avviker från huvudöversättningen.

Inte heller när det gäller fraserna är det svårt att stryka en del, återigen på grund av likheterna mellan svenska och nederländska. Ordet *stå* t.ex. kan bl.a. betyda *finnas skriven*, vilket utan problem kan översättas med det nederländska *staan*. I engelskan finns det emellertid inte någon täckande översättning. Därför har man tagit upp en rad fraser som illustrerar de olika sätt på vilka man kan översätta det svenska *stå* till engelska. I alla dessa fraser kan man använda ordet *staan* på nederländska. I det här fallet har vi strukit alla fraser med undantag av en, som har som funktion att visa för den svenska användaren att han/hon kan använda ordet *staan* om *stå* betyder *finnas skriven*.

Sammanfattningsvis kan man säga att det är vår erfarenhet att det finns färre översättningsproblem mellan svenska och nederländska än mellan svenska och engelska. Därför åstadkoms minskningen med 25 procent mer eller mindre automatiskt och utan stora problem, nämligen som följd av att vi omarbetar en svensk-engelsk ordbok till en svensk-nederländsk.

#### 4. En annan ordbok

Minskningen av omfånget är en av orsakerna till att vår ordbok kommer att se helt annorlunda ut än den svensk-engelska. Så pass mycket att jag tror att många användare inte eller knappast skulle se att vi grundar oss på *Stora svensk-engelska ordboken*. Det finns ytterligare två skäl för det.

För det första har vår ordbok en helt annan mikrostruktur. En ordboksartikel indelas i tre delar: '*huvud*', '*nollsteg*' och '*fraseologi*'. I artikels huvud återges uppslagsordet, följt

av ordets ordklassstillhörighet och böjning. Om uttalet avviker från de gängse reglerna ges också uttalet.

Nollsteget omfattar alla översättningar. Huvudöversättningar förses med koder: 0.1, 0.2 etc. Om det finns mer än en översättning i den svensk-nederländska delen, ska redaktören förklara för de svenska användarna i vilket avseende översättningarna skiljer sig från varandra i betydelse, eftersom användaren annars inte vet vilken översättning han/hon ska välja. I de flesta fall sker det via en kort sammanfattning på svenska eller en etikett för ämnesområde eller stilnivå.

I fraseologin, slutligen, ges exempel, idiom, fasta förbindelser och grammatiska kollokationer (dvs. en kombination av ett ord med en preposition). Varje fras får en kod som består av två siffror. Den första anger ordklassstillhörigheten hos det ord som uppslagsordet kombineras med. Den andra siffran hänvisar till en av de (kodade) översättningarna. Frasen *vidta åtgärder* som behandlas under uppslagsordet *åtgärd* får t.ex. koden 3.1. Trean betyder att ordet *åtgärd* kombineras med ett verb (*vidta*) och ettan betyder att det gäller den betydelse av *åtgärd* som man kan hitta i nollsteget och som behandlas under koden 0.1.

Figur 1. Uppslagsordet 'hem' i svensk-nederländska ordboken

**hem'** I <s;~met,=> 0.1 [bostad] *huis, woning* → <hushåll> *huishouden* 0.2  
[institution för vård] *tehuis* ♦ 2.1 komma från ett gott ~ *van goede*  
*huize komen* 3.1 sköta ~met *het huishouden doen*  
II <adv> 0.1 *naar huis*

För det andra lägger vi i några fall till information som inte finns i den svensk-engelska ordboken. *Stora svensk-engelska ordboken* är en ordbok för aktiv användning, dvs. den har som viktigaste målgrupp svenskspråkiga som vill producera text på engelska. Den svensk-nederländska delen av vår ordbok är emellertid både en aktiv ordbok för svensktalande (som vill producera text på nederländska) och en passiv ordbok för nederländsktalande (som vill förstå svensk text eller som vill slå upp ett svenskt ords uttal eller böjning). Eftersom uttal och böjning inte finns i den svensk-engelska ordboken, måste vi lägga till den informationen.

Ibland är det nödvändigt att ge viss information två gånger, en gång på svenska och en gång på nederländska. Om en svensk användare får informationen att en viss översättning även kan användas bildligt genom tillägget <även bildligt>, så har den nederländskspråkiga användare rätt till samma information: <ook figuurlijk>. Det är självklart att redaktörerna försöker att undvika sådana dubbla informationer. Om man t.ex. vill ange att både det svenska ordet *mus* och det nederländska ordet *muis* också kan betyda den lilla apparaten vid datorn, kan man i sig förse översättningen med två kommentarer:

**mus** <s;~en,möss> 0.1 <äv. data.> **muis** <ook comp.>

Nackdelen är emellertid att artikeln blir mycket översködlig och att man nästan glömmer att *mus/muis* i första hand är ett djur. I detta fall kan man i stället skriva:

**mus** <s;~en,möss> 0.1 <zool.> **muis** 0.2 <data.> **muis**

När det gäller etiketterna har vi strävat efter att hitta förkortningar som man känner igen oberoende av vilket modersmål man har. Så använder vi förkortningen <mat.> för *matematik*,



även om det vanliga nederländska förkortningen är <wisk.> (*wiskunde*). Ordet *mathematiek* förekommer dock också i nederländskan, fast det inte används så ofta.

Den dubbla uppläggningsen av ordboken verkar vara mycket ambitiös och man skulle kunna frukta att ordboken inte blir målinriktad. Med så (relativt) små språkområden som nederländska och svenska kan man emellertid inte välja en viss målgrupp, såsom man kan göra när det gäller världsspråk (engelska, franska o.d.) eller kombinationer av ett världsspråk och ett litet språk (engelska och svenska, engelska och nederländska etc.). I vårt fall måste man helt enkelt rikta sig till alla som på något sätt sysslar med de båda språken.

Dessutom skulle jag inte vilja gå så långt som att kalla en sådan uppläggning för en omöjlighet. Jag ser det framför allt som ett praktiskt problem: Det är svårt för en nederländsk-språkig redaktör att bedöma vilka problem en svensk användare kan ha med att uppfatta den information som ges i en ordboksartikel. Det är bl.a. därför som vårt material granskas mycket noggrant av både nederländskspråkiga och svenskspråkiga.

## Noter

1. Nu då Sverige har ansökt om medlemskap i EG kan man anta att behovet av en ordbok bara blir större.
2. Van Dale är Nederländernas ledande ordboks förlag som bl.a. ger ut en stor nederländsk ordbok i tre delar; dessutom har förlaget publicerat stora tvåspråkiga ordböcker (engelska, franska och tyska). Den svensk-nederländska/nederländsk-svenska ordboken ska ingå i serien handordböcker (nederländska, franska, engelska, tyska, spanska).
3. Se konferensbidraget av R. Schutz (Van Dale Lexicografie).

## Litteratur

- Boer-den Hoed, P.M. 1956. *Zweeds handwoordenboek*. Den Haag.  
 Elsevier, 1984. *Woordenboek Zweeds-Nederlands/Nederlands-Zweeds*. Amsterdam, Brussel.  
 Esselte, 1988. *Stora svensk-engelska ordboken (A Comprehensive Swedish-English Dictionary)*. Stockholm.  
 UbU18, 1989/90. *Utbildningsutskottets betänkande*. Stockholm.  
 Wolters, 1988. *Wolters Miniwoordenboek Zweeds-Nederlands/Nederlands-Zweeds*. Groningen.

## Den selektive specialordbog

Den selektive specialordbog er en oversættelsesordbog, der fokuserer på underområder inden for fagsproglige enkeltfag. I kraft af det selektive tilsnit, kan ordbogen både være slank og alligevel samtidig satse på høj kvalitet. Makro- og mikrostruktur kan således med fordel indrettes specifikt for hver enkelt ordbog. For at kunne arbejde med fremstilling af sådanne ordbøger som undervisningsdisciplin, er det nødvendigt at have meget klare kriterier for det selektive emnes afgrænsning. Disse problemer diskuteres i lyset af to selektive ordbøger, hver med sin specifikke struktur.

Indledningsvis skal det fastslås, at min primære interesse i ordbogsarbejde er at dyrke det som et emne ved den fagsproglige oversættelsesuddannelse på højeste niveau ved Handelshøjskolen i København. Jeg har derfor det dobbelte sigte at udvikle et koncept, der kan tilgodese emnet som en undervisningsdisciplin og samtidig tilgodese det som et produkt, der kan markedsføres.

### Den selektive model

Hvis det skal lykkes at producere dansk-engelske og engelsk-danske fagsproglige oversættelsesordbøger på overskuelig tid med studerende som leksikografer, er det indlysende nødvendigt, at der arbejdes med en meget slank ordbogsmodel.

Samtidig er det af hensyn til de studerende vigtigt, at ordbogsarbejdet får høj kvalitet, men dette er også vigtigt for markedsføringen af det færdige ordbogsprodukt. Fagsprogsmarkedet har kun interesse i slanke ordbøger, hvis de udgør et vægtigt bidrag til eksisterende ordbøger.

Betegnelsen selektiv er valgt, fordi valget af lemmata er stærkt selektivt ved kun at fokusere på underområder inden for fagsproglige enkeltfag (Nielsen 1990:32ff). Den selektive ordbog har dermed en anden indfaldsvinkel end dem, der omfattes af betegnelsen minimerende fagsproglige ordbøger (Hausmann 1977:5f og Nielsen 1990:32ff), idet disse forfattere med minimerende mener ordbøger med et overfladisk udvalg fra et fagområde.

### To eksempler

I det følgende skal omtales to ordbøger, hvor afgrænsningen rent praktisk har været foretaget ud fra helt forskellige kriterier.

Den ene ordbog, *Joint Venture - engelske og danske termer*, (Lenstrup 1989b) fokuserer på et underområde af emnet virksomhedsformer, som kan være aktieselskaber,

kooperativer, etc., et emne hvis mere traditionelle betegnelser på de forskellige delområder kan slås op i eksisterende ordbøger.

De indgangsord, der indgår i ovennævnte ordbog, er imidlertid selekteret fra en stor mængde faglitteratur ud fra det kriterium, at de skal supplere eksisterende oversættelsesordbøgers dækning af fagområdet. Grænsen, for hvad der skal indlemmes, og hvad der ikke skal, er dog ikke blot et mekanisk spørgsmål, om hvorvidt en fagterm kan slås op andetsteds. Men det er et spørgsmål, om hvorvidt det pågældende lemma har betydningsnuancer, som måtte være specifikke for vores ordbogs emne, og som ikke fremgår i de gængse fagordbøger.

Den anden ordbog (forventet udgivelse primo 1992) er selektiv efter et andet og mere praktisk princip. Ordbogens emne er indkomstskat, som ikke er dækket af nogen eksisterende tosprogsordbog med engelsk og dansk, hvorfor en omfattende ordbog på området er et oplagt emne. Men på grund af de begrænsede produktionsressourcer er der truffet et indsnævrende valg, nemlig selvangivelsen, der indsnævredes yderligere til den engelske selvangivelse - the tax-return.

Denne form for afgrænsning, der bestemmes af formularen med tilhørende vejledning, er naturligvis langt lettere at håndtere end den ovenfor omtalte. Og med studerende som leksikografer er den mere hensigtsmæssig, netop fordi den fritager leksikografen for overvejelser, om hvad der har relevans som lemmata.

Sådanne overvejelser kan nemlig være vanskelige for studerende, der ikke kan forventes at have fornøden kompetence på et fagsprogligt område til at afgøre, hvad der er specifikt for et enkeltfags underområder, og hvad der hører til fagets overordnede begreber. Efter min erfaring fører denne usikkerhed til spild af tid og til frustrationer, der forringer koncentrationen om det konstruktive ordbogsarbejde for de studerende.

## Struktur

Med hensyn til makrostrukturen forestillede vi os ved starten af arbejdet ved begge ordbøger en traditionel model med to dele, en for hvert sprog, hver med sin alfabetiske indgang, som vi kunne opbygge sideløbende. Men i begge tilfælde viste en sådan fremgangsmåde sig at være uhensigtsmæssig, nemlig af årsager, der hang sammen med karakteren af de områder, vi arbejdede med.

Ved emnet skat drejede arbejdet sig om at finde ækvivalenter til termer i en kulturbundet institution, som ganske vist begge lande har, idet der betales indkomstskat, arveafgift, o.s.v. Men de traditioner og strukturer, der ligger bag, er så forskellige, at det viste sig mest praktisk at tage hvert land for sig, altså starte med det engelske, finde forklaringer og danske ækvivalenter og udgive det som én ordbog. Og så på et senere tidspunkt gå den modsatte vej og udgive en dansk engelsk ordbog om den danske selvangivelse.

Enhver forestilling om at alfabetisere de danske ækvivalenter fra en sådan engelsk-dansk ordbog, eksempelvis om skat, med de engelske indgangsord koblet på til en dansk-engelsk ordbogsdel — nemlig det der sker ved visse oversættelsesgenererede fagsprogsordbøger — bør man droppe. En sådan dansk-engelsk ordbog ville være et kunstprodukt, og det ville ikke være venligt at friste den travle bruger til at tro, at en alfabetiseret dansk del med tilkoblede engelske indgangsord fra den engelske del, kunne anvendes ved oversættelse fra dansk til engelsk om *danske* skatteforhold.

Emnet 'joint venture' er derimod et internationalt emne, og langt den meste tilgængelige litteratur er på engelsk. De problemer, der behandles i danske tekster om joint

venture, er kendte og omtalte problemer i den engelske sprogverden. Her er altså ikke tale om to kulturelt forskellige verdener, her er tale om den globale verden. Og her bunder anisomorfien i det forhold, at dansk er et lille sprog og en lille verden, og engelsk er et stort sprog og en stor verden, og emnet joint venture er et internationalt emne.

Dette forhold har vi ladet få indflydelse på makrostrukturen. Der er således kun ét alfabetisk indeks i ordbogen, der er bygget op omkring de engelske begreber, men der er både engelske og danske opslagsord i dette indeks.

I mikrostrukturen er såvel alle opslagsord som alle ækvivalenter illustreret med kontekster, og alle kontekster har kildeangivelse. Kildeangivelsen tjener til at give brugeren oplysning om kildens årstal og autoritet, hvor især det sidste er vigtigt for den fagsproglige oversætter, hvis han skal oversætte ikke-etablerede betegnelser (Fabritius 1991:12). Endelig giver kildeangivelse naturligvis mulighed for at søge litteratur om emnet. Makro- og mikrostruktur er i øvrigt beskrevet i Lenstrup (1989a:16f) og Lenstrup (1990).

### Grænseområder til almensproget

Jeg går ud fra som givet, at en seriøs fagsproglig ordbog tager udgangspunkt i faget, og at brede fagsproglige ordbøger i forhold til almensproglige ordbøger medtager det sprogområde, der ligger på overgangen mellem faget og almensproget, dér hvor fagmand og lægmand kommunikerer med hinanden, f.eks. betegnelser for sygdomme, som ikke er latinske betegnelser, osv.

Dette sprogområde udgøres i stor udstrækning af veletablerede gloser, som uden videre fungerer i kommunikationen. Men den selektive specialordbog gør ingen eller meget få udflugter til dette sprogområde. Den tager sig derimod af nye og vanskelige størrelser, der så efterhånden etableres eller glider ud af brug. Når nye betegnelser er etableret, behøver de ikke længere den dokumentation, som er nødvendig ved første præsentation.

Som et eksempel på en glose, der i de senere år er gledet helt over i almensproget og nu kan indlemmes dér, kan nævnes **netværk**, der med netværkstankens udbredelse er gledet fra specialområderne som f.eks. **økonomisk netværk: firmagrupper** til at betyde mere eller mindre organiserede kontakter og samarbejder.

På det skatteretlige område og især det, der har betydning for oversættelse af skatteretlige dokumenter for klienter, nemlig det område, der vedrører undgåelse af dobbeltbeskatning, er hjemstedskriteriet et vigtigt spørgsmål, hvilket bl.a. fremgår af den grundige behandling, spørgsmålet får i overenskomsten mellem Storbritannien og Danmark til undgåelse af dobbeltbeskatning. Der anvendes dér stor omhu på at definere, og på at definere de begreber, der defineres med.

Det er indlysende, at det er af interesse for skatteyderen at undgå dobbeltbeskatning, men det er af lige så stor interesse for myndighederne at forhindre skatteunddragelse. Parterne i sådanne sager kan derfor have modstridende interesser i fortolkningen af hjemstedsbegrebet. Og translatøren kan risikere alvorlige konsekvenser for sin klient, hvis han vælger en uheldig ækvivalent fra en ordbog.

Hvor meget skal skatteordbogen have med på et sådant punkt? Skal den citere love, cirkulærer og lignende fra to eller flere lande til støtte for oversætteren i hans vurdering, af hvorledes noget så enkelt som bopæl skal oversættes?

Det mener jeg ikke, den absolut skal, men den kan anvise ækvivalenter ved citater fra de foreliggende paralleltekster, f.eks. dobbeltbeskatningsoverenskomsterne, suppleret med

henvisninger til relevante steder i skattelovgivningen, skattemyndighedernes vejledninger, revisionsfirmaernes fortolkninger, skatteretlige domstoles afgørelser, etc.

Overvejelserne, om hvad der skal med i denne type ordbog, kan således blive ganske forskellige fra overvejelserne, om hvad der skal med i en mere generel fagsprogsordbog. Sideløbende hermed sker der det, at den stadige præcisering af f.eks. skattemæssige begreber inden for OECD-landene og harmoniseringen inden for EF medfører, at den selektive specialordbogs område hurtigt ændres. Noget kan glide over i mere generelle fagsprogsordbøger, andet er måske forældet og kan glide ud.

### Specialordbøger kontra flerfagsordbøger

Ideen om at snævre en fagsproglig oversættelsesordbog ind til et lille afgrænset område fremføres også af Nielsen (1990:36), der hævder "...at det leksikografisk set er bedre at udarbejde enkeltfagsordbøger end flerfagsordbøger, da enkeltfagsordbøger er ideelle til at give en dybtgående og detaljeret leksikografisk behandling af et i øvrigt vanskeligt tilgængeligt fagområde."

Problemerne ved ikke at foretage en sådan indsnævring fremgår hos Dyrberg og Tournay (1990:273) i de afsluttende bemærkninger i en artikel om principperne for en fransk-dansk fagsproglig oversættelsesordbog inden for "områderne økonomiske, juridiske og politiske forhold samt forvaltningen."

Dyrberg og Tournay skriver dér, idet jeg citerer fra en dansk version: "at dét der, foruden fastlæggelsen af ækivalens, har voldt os de største problemer og fortsat volder problemer, er en konsekvent implementering af vore principper. Disse vanskeligheder har deres rødder i de meget forskelligartede strukturer i de områder, vi behandler."

Hermed sigter forfatterne primært til anisomorfi på de to sprogs parallelle fagområder, men hvis "forskelligartede strukturer" er en vanskelighed dér, så multipliceres vanskeligheden med antallet af fagområder i ordbogen.

Løsningen på dette problem kunne naturligvis være at dele ordbogen op i mindre bind. Det forudsætter, at målgruppen kan skelne mellem områderne og ved hvor underområderne hører til. Men forfatterne har angivet deres målgruppe til at være "ældre franskstuderende samt oversættere (og erhvervsfolk)", og på det niveau kan brugerne udmærket skelne virksomhedsøkonomi fra nationaløkonomi og forvaltningsret fra politik.

Opdeling i specialordbøger giver mulighed for emnetilpassede principper for makro- og mikrostruktur, og selvfølgelig også mulighed for at revidere enkeltområder efter behov. Og dermed et ordbogsprodukt, der foruden at være hensigtsmæssigt for den professionelle ordbogsproducent også måtte være hensigtsmæssigt for brugerne samt for lærere og studerende, når de optræder som nicheleksikografer.

### Litteratur

- Fabritius, Jeanny. 1991. TermData - PC-Termregistrering og -oversættelse i operationel form. *CLM-nyhedsbrev*, Cand.ling.merc.foreningen, nr. 3, 1991, København, (12).
- Dyrberg, Gunhild og Joan Tournay. 1990. Définition des Équivalents de Traduction de Termes Économiques et Juridiques sur la Base de Textes Parallèles. *Cahier de Lexicologie* no.56-57, publiés par Bernard Quemada, Didier Erudition, 1990, 1 et 2, 261-274. Paris.

- Hausmann, Franz Josef. 1977. *Einführung in die Benutzung der neufranzösischen Wörterbücher*. Niemeyer, Tübingen.
- Lenstrup, Rita. 1989a. ORDBØGER - "De er så mange slagse". *SPRINT* 1989-2 Sproginstitutternes Tidsskrift, Handelshøjskolen i København, 13-20.
- Lenstrup, Rita. 1989b: *Joint venture - engelske og danske termer*. Samfundslitteratur, København.
- Lenstrup, Rita. 1990: Compilation of bilingual specialized LSP dictionaries. Indsendt til publicering i *Proceedings from The 7th European Symposium on Language for Special Purposes*, Budapest 1989. Kopi af manus kan rekvireres hos forfatteren.
- Nielsen, Sandro. 1990. Noget om fagsproglige ordbøger og fag. *Netværk LSP* 2, juni, Handelshøjskolen i Århus, 31-38.

Frede Boje og Anna Braasch

## Hvad får man skudt i skoene? Flerordsenheder i aktive ordbøger for mennesker og maskiner

Det arbejde, vi her fremlægger, er udsprunget af et samarbejde inden for den danske EUROTRA-gruppe, men har udviklet sig til et samarbejde mellem artiklens to forfattere som repræsentanter for hhv. EUROTRA-DK (som nu er en del af det nyoprettede Center for Sprogteknologi) og en projektgruppe ved Handelshøjskolen i København. Formålet med det igangværende arbejde er at samle og klassificere et så omfattende materiale af eksempler på flerordsenheder, at vi vil kunne registrere de mønstre, der forhåbentlig vil tegne sig, og på grundlag heraf udvikle en beskrivelsesmodel, som er eksplicit og udtømmende nok til både at være implementerbar i NLP og kunne danne basis for en aktiv ordbog for mennesker. Dette materiale hentes fortrinsvis i maskinlæsbare korpura som DK87-90 og sammenholdes med det materiale, som er beskrevet i Erik Bruun: Dansk Sprogbrug. Metoden og målsætningen omfatter både leksikografiske og datamatiske aspekter, og udbyttet af vekselvirkningen imellem disse aspekter anser vi for at være det mest spændende i dette projekt.

### 0. Indledning: To indfaldsvinkler: human- og maskinoversættelse

Når ovenstående udtryk er valgt som titel, er det for at illustrere den ene af de to indfaldsvinkler til vores arbejde; det er nemlig et eksempel på et konkret oversættelsesproblem, vi er stødt på i forbindelse med undervisningen af danskere i oversættelse fra dansk til tysk. På dansk har vi udtrykket:

at skyde nogen noget i skoene

og på tysk udtrykket:

jemandem etwas in die Schuhe schieben.

Det er nærliggende at tro, at de to udtryk betyder det samme og bruges på samme måde, og den opfattelse får man bestyrket, hvis man slår op i Gyldendals røde ordbøger Dansk-tysk og Tysk-dansk:

I. *sko*: Schuh, [...]; skyde én ngt i -ene: e-m etw in die Schue [sic!] schieben, e-m etw unter' stellen; [...]

(Bork 1990a)

*Schuh* m -e. 1. sko, [hoher] ~ støvle (e-m etw in die -e schieben); [...]

(Bork 1990b)

Vores erfaring siger os imidlertid, at der er en forskel i anvendelsen af udtrykket på dansk og tysk, idet de leksikalske og syntaktiske restriktioner for objektet er forskellige, men problemet er: Hvor finder man tilstrækkelig præcise oplysninger om, hvordan udtrykket bruges på dansk og/eller tysk?

Der er os bekendt ikke nogen, som har foretaget omfattende og systematiske undersøgelser af, hvorledes den slags udtryk er behandlet i danske ordbøger, mens der er i hvert fald to, som - næsten samtidig, men uafhængigt af hinanden - har undersøgt tyske ordbøgers behandling af fraseologiske udtryk. Det drejer sig om Anne Lise Kjær og Anna Braasch, som begge først har skrevet en specialeafhandling (Kjær 1986) og (Braasch 1987) og dernæst en artikel, som sammenfatter afhandlingens resultater (Kjær 1987) og (Braasch 1988).

Ikke uventet konstaterer man, at de eksisterende opslagsværker ikke giver tilfredsstillende svar på vores spørgsmål om restriktionerne for objektet<sup>1</sup>.

I vores korpusundersøgelser har vi fundet meget få eksempler på brugen af udtrykket i overskriften, men sammen med de oplysninger, som kan stykkes sammen fra de forskellige ordbøger, underbygger de vores opfattelse: Mens objektet på tysk normalt er substantivet "die Schuld"<sup>2</sup>, føles "skylden" som objekt på dansk som "dobbelts konfekt"<sup>3</sup>. På dansk er objektet det, man får skylden for - som regel, men ikke nødvendigvis, udtrykt ved en at-sætning<sup>4</sup>, hvilket tilsyneladende ikke er muligt på tysk.

Oversættelsesproblemer som dette, som især dukker op ved oversættelse fra modersmål til fremmedsprog (men ikke kun dér), har fået os til at stille spørgsmålet: Hvordan udarbejder man en leksikalsk beskrivelse, som er detaljeret og præcis nok til at give svar på den slags spørgsmål?

Vores anden indfaldsvinkel vil vi i det følgende give en noget grundigere redegørelse for. Det drejer sig om dét behov for en detaljeret og præcis beskrivelse af leksikalske enheder, man har i arbejdet med maskinoversættelse. Men først vil vi uddybe, hvad det er for udtryk, vi ønsker at beskrive (afsnit 1), og kort omtale de særlige krav, maskinoversættelsen stiller til ordbøger og grammatikker (afsnit 2).

## 1. Definition af flerordsenheder - kriterier - typologi

Desværre bruger forskellige forfattere et utal af forskellige betegnelser om de fænomener, vi undersøger, samtidig med, at mange af betegnelserne, fx "kollokation", bruges med forskelligt indhold. Pilz (1978:8) klager over "diese geradezu chaotische terminologische Vielfalt", og situationen synes ikke at være blevet bedre siden da<sup>5</sup>.

Det, vi ønsker at beskrive, er, hvad vi - som et forhåbentlig neutralt overbegreb - vil kalde *flerordsenheder*. Ved en flerordsenhed forstår vi:

En sproglig størrelse, som består af mere end ét ord, og hvis betydning og/eller brug ikke kan beskrives udtømmende alene ved generelle syntaktiske regler og en leksikalsk beskrivelse af enkeltordene.

Når vi nedenfor giver et forslag til, hvorledes man kan inndele flerordsenhederne, betyder det ikke, at vi har et gennearbejdet system byggende på klare definitioner, men at vi anser det for nødvendigt at skitsere, hvordan de forskellige fænomener efter vores mening er relateret. Vores hovedinteresse ligger imidlertid ikke i denne "grov"-typologi, men, som det vil fremgå



af det følgende, i en systematisk klassifikation af de forskellige typer kollokationer og fraseoleksemer.

For brugeren - inklusive edb-maskinen - er det underordnet, om et givet udtryk kaldes en kollokation, et fraseoleksem eller noget tredje. Det afgørende for ham (eller den) er, at fænomenet er registreret og beskrevet på en måde, som sætter brugeren - og altså også maskinen - i stand til at vælge den korrekte form (evt. en af de korrekte former). Med udtrykket "at vælge" tænker vi her både på receptionsfasen - inklusive den maskinelle analyse - og produktionsfasen.

Ved arbejdet med registreringen og klassifikationen af udtrykkene er det efter vores mening hensigtsmæssigt at bruge størrelse og stabilitet som hovedkriterier. Men ved en udtømmende teoretisk *beskrivelse* af flerordsenhederne er der mange flere faktorer, der skal inddrages. En sådan beskrivelse ville imidlertid fuldstændig sprænge rammerne for denne artikel. Vi vil derfor begrænse os til at nævne de vigtigste faktorer i stikordsform og kort kommentere dem. Det er oplagt, at der er en nær sammenhæng og indbyrdes afhængighed mellem flere af faktorerne.

- 1 Størrelse og ordklassetilørsforhold: 2 ord - sætningsled - sætning - periode - flere perioder.
- 2 Stabilitet: fra fuldstændig fri kombination til helt fast forbindelse.
- 3 Metaforik: konkret - metaforisk.
- 4 Opacitet: transparent - opak.
- 5 Kompositionalitet: kompositionel - ikke-kompositionel.
- 6 Hyppighed: højfrekvent - lavfrekvent kombination.
- 7 Normbundethed: afhængighed af sproginterne og/eller sprogeksterne normer.
- 8 Oversættelighed.
- 9 Forudsigelighed.

Hertil kommer et par faktorer, som ikke er væsentlige for, hvordan de forskellige typer flerordsenheder defineres, men som til gengæld har betydning for, hvilke af typerne man især vil beskæftige sig med, nemlig:

- 10 Teksttyper.
- 11 Emneområder.

Ad 1. Pr. definition må en flerordsenhed bestå af mindst to ord, mens definitionen ikke sætter nogen fast grænse opad. Da vi ønsker at registrere udtryk, som det vil være relevant at have som ordbogsindgange i en maskinlæsbar ordbog - og det vil i praksis stort set svare til lemmaer og sublemmaer i en almindelig ordbog - interesserer vi os især for størrelser, der tilsammen enten udgør ét sætningsled, typisk et NP uanset ledfunktion, eller danner kernen i en verbalfrase, typisk verbum + objekt(er). Derfor har vi heller ikke foretaget en underinddeling af længere idiomatiske udtryk, som utvivlsomt er relevant i anden sammenhæng<sup>6</sup>.

Ad 2. Stabilitet. De to yderpunkter er ret uproblematisk. Til gengæld er det beskrivelsen af de fænomener, som ligger derimellem, der især er interessant. Er de restriktioner, der gælder for variationsmulighederne for de enkelte elementer i flerordsenhederne morfologiske eller syntaktiske? Er de bundet til en lukket liste af ord eller til klasser af ord med fælles semantiske egenskaber? Er der en afhængighedsrelation mellem udfyldningen af

forskellige pladser i et komplekst udtryk? Det er alle disse faktorer, vi først og fremmest vil belyse, monolingvalt og kontrastivt.

Ad 3. Metaforik. Mange flerordsenheder er mere eller mindre levende metaforer, og den rutinerede oversætter ved, at metaforer ofte giver særlige oversættelsesproblemer, men metaforik er næppe egnet som inddelingskriterium.

Ad 4-5. Opacitet og kompositionalitet er to betegnelser, der - med forskellig synsvinkel - bruges til at beskrive det samme fænomen. Samtidig har de nær relation til pkt. 9, forudsigelighed. Der skelnes gerne mellem forskellige grader af opacitet, mens kompositionalitet/ikke-kompositionalitet oftest ses som en dikotomi. En kompositional konstruktion vil være transparent, en ikke-kompositional mere eller mindre opak. Begge begreber står i relation til afkodning (reception). Selv om kompositionalitetsbegrebet er problematisk, mener vi, at det er rimeligt at bruge det som kriterium for, om en leksikalisk betinget flerordsenhed er et idiomatisk udtryk eller en kollokation.

Ad 6. Hyppigheden af en kombination anses af nogle for et kriterium for, om den er en kollokation, se fx Bergenholtz (1989:777). Efter vores mening er høj forekomstfrekvens et stærkt indicium for, at en ordkombination er en flerordsenhed, men ikke anvendeligt som kriterium.

Ad 7. Normbundethed. Vi mener, at den tredeling i "system - norm - sprogbrug", som Coseriu har indført og v. Polenz har videreudviklet<sup>7</sup>, er bedre egnet til at forklare nogle af de mekanismer, der er på spil, end den Saussure'ske todeling i langue - parole, men tror ikke, at det er nødvendigt at inddrage normbegrebet som klassifikationskriterium. I sin afhandling om "Normbetingede ordforbindelser i tysk juridisk sprog" (Kjær 1990a) har Anne Lise Kjær påvist, at retsvirkninger kan være afhængige af, at en sags parter overholder præskriptive normer for formuleringer, men så vidt vi kan skønne, er der her tale om noget, som især (og måske kun) er relevant for juridisk sprogbrug, så vi vil tillade os i første omgang at lade problemet ligge.

Ad 8-9. Forudsigelighed og oversættelighed hører til en vis grad sammen. Hvis et givet syntagma er fuldstændig forudsigeligt, altså hvis det kan konstrueres korrekt alene ud fra kendskab til generelle syntaktiske regler og enkeltordenes betydning, skulle udtrykket ud fra den ovenfor nævnte definition falde uden for emnet. Problemet er, at det ofte først er i forbindelse med oversættelse, at man opdager, at der er nogle restriktioner. Hausmann (1989:1013) sætter sagen på spidsen ved at sige, at hvis man sammenligner ækvivalerende udtryk på alle verdens sprog, vil resultatet nok blive, at et hvilket som helst udtryk, som formodes at være en fri forbindelse, viser sig at være en kollokation. Selv om det selvfølgelig er en vild overdrivelse, er der mere om snakken, end folk, der ikke arbejder kontrastivt, i almindelighed tror.

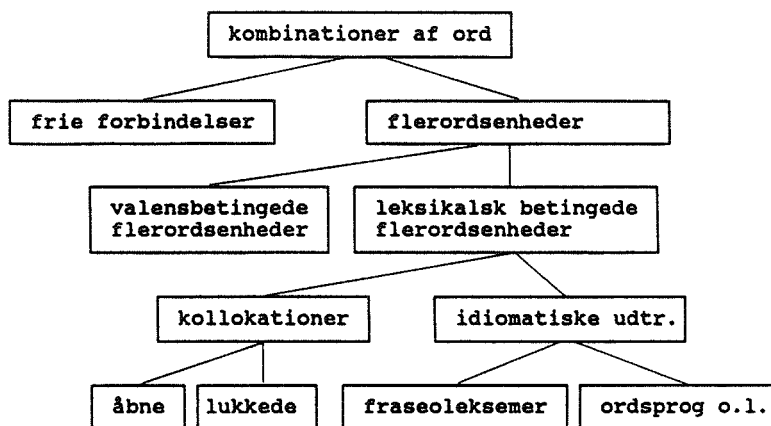
I en redegørelse for sit arbejde med Grunddeutsch nævner Pfeffer (1979:311f), at i ca. 80 % af korpusbelæggene for verbet *halten* indgår det i "redensartige Bindungen", og at hverdags sproget er mere end 15 % "redensartig gebunden". Selv om disse oplysninger skal tages med et vist forbehold, antyder de problemets omfang, specielt for de to former for oversættelse, vi fokuserer på: maskinoversættelse og oversættelse til fremmedsprog.

Forudsigelighed er et begreb, som er temmelig umuligt at operere med, i hvert fald hvis man ikke begrænser det til kun at være forudsigelighed i forbindelse med oversættelse, og i så fald skal man tage stilling til, hvilken type oversættelse mellem hvilke sprog.

Konklusionen på dette er, at selv om netop den manglende forudsigelighed er en af hovedårsagerne til, at der er behov for en fyldestgørende beskrivelse af flerordsenheder, er forudsigelighed uanvendelig som egentligt klassifikationskriterium.

Ad 10-11. På grund af formålet med vores projekter er det klart, at vi ikke beskæftiger os særligt med fænomener, som udelukkende eller helt overvejende optræder i talesprog eller i skønlitteratur. Vi vil her henvise til Skyum-Nielsen (1991), som ud fra en anden indfaldsvinkel behandler mange af disse fænomener. Vi lægger hovedvægten på, hvad der er vigtigt i fagsproglige tekster samt andre non-fiktive tekster, altså fx avisartikler. Det indebærer bl.a., at forbindelser, som er okkasionelle, altså ad hoc-dannelser, og anden kreativ sprogbrug principielt ikke inddrages<sup>8</sup>.

Den typologi, som er det foreløbige resultat af vores overvejelser, er vist i fig. 1:



Figur 1

Skemaet er bl.a. inspireret af Sonja Poulsens og Peder Skyum-Nielsens arbejder<sup>9</sup>, men er mere detaljeret end Poulsen og mindre detaljeret (og principielt anderledes) end Skyum-Nielsen. Vi må igen her afstå fra en fuldt udbygget argumentation og blot kort kommentere de vigtigste distinktioner.

Det, vi her har kaldt "valensbetingede flerordsenheder", svarer stort set til, hvad BBI (1986) kalder "grammatical collocations". Når vi skiller dem ud som en særlig gruppe, skyldes det ikke kun, at de af de fleste forskere med god grund ikke regnes med til kollokationerne, men også, at det er fænomener, som kan beskrives eksakt og udtømmende med allerede kendte metoder, også til brug ved maskinel analyse og oversættelse. Inden for EUROTRA er der udviklet et meget fintmasket system til valensbeskrivelse. Systemet er implementeret i ordbøger og grammatikker, og det fungerer.

Kriteriet for distinktionen mellem kollokationer og idiomatiske udtryk er idiomaticiteten: kollokationerne er ganske vist uforudsigelige, men kan forstås kompositionelt, mens de idiomatiske udtryk hverken i produktion eller reception er kompositionelle. Fraseoleksemerne er (som de almindelige leksemer) "byggeklodser" i sætninger, men aldrig hele sætninger.

## 2. Aktive ordbøger for mennesker og maskiner

Ordbøger for mennesker, både et- og tosprogede, aktive og passive udnytter (omend på forskellig måde og i forskellig grad) brugerens evne til at

- \* generalisere ud fra givne sproglige eksempler og dermed danne analogier;
- \* udlede oplysninger der er implicit givne i eksempler og dermed få adgang til 'skjulte' regulariteter eller uregelmæssigheder;
- \* overføre sin viden om sprogs udtryksmuligheder fra modersmålet til målsproget i form af sprogkompetence og intuition;
- \* inddrage sin almenviden om verden.

I automatisk behandling af naturlige sprog (Natural Language Processing, NLP), herunder maskinoversættelse (Machine Translation, MT), kan man ikke bygge på disse evner. Dette betyder, at maskinen skal have kompenseret for det manglende vidensgrundlag med eksplicite oplysninger i en eller anden form. Derfor vil en ordbog til NLP altid indeholde flere og mere detaljerede oplysninger, også om kildesprogets enheder, end en aktiv ordbog for mennesker ved samme enhed (opslagsord).

### 2.1 EUROTRAs ordbogssystem

EUOTRAs oversættelses- og dermed også dets ordbogssystem bygger på princippet om multilingvalitet. Dette indebærer på det overordnede plan, at der for hvert sprog udarbejdes en monolingval ordbogs-komponent. Vi vil her ikke gå nærmere ind på, hvordan denne komponent er opbygget, blot nævne, at systemet fordeler oversættelsesprocessen på flere niveauer. Hvert niveau har sin egen ordbog, og der er en "oversættelses"ordbog mellem hvert niveaupar. For overskuelighedens skyld omtales i det følgende kun grænsefladeordbøgerne, der indholdsmæssigt mest ligner aktive ordbøger for mennesker<sup>10</sup>.

Den egentlige oversættelsesordbog, den såkaldte transferordbog, knytter ideelt kun ordbogsindgangene fra to monolingvale grænsefladeordbøger sammen. Dette betyder, at de monolingvale MT-ordbogs-komponenter stiller større krav til detaljeringsgrad og systematik end ordbøger for mennesker.

Ved analysen af en tekst, inden den kan oversættes, skal også flerordsenhederne identificeres, deres komponenter samles til en leksikalsk enhed; dette er særlig vigtigt for sådanne flerordsenheder, hvis betydning ikke kan forstås ud fra komponenternes.

Principielt skulle de monolingvale ordbøger kunne kombineres på kryds og tværs i EUROTRA-systemet. Der er dog en del væsentlige problemer, som endnu er uløste, det gælder især på området flerordsenheder.

### 2.2 Grammatik og ordbog i MT

Rollefordelingen mellem grammatik og leksikon (dvs. ordbog) er nøje defineret i MT-systemer (og for NLP i almindelighed). Grammatikken beskriver det generelle (regler for ordstilling, bøjningsmønstre...), leksikonet beskriver det individuelle, ordenes egenskaber hver for sig (ordets betydning(er), rektion, kompatibilitet med andre ord, genus, bøjning...). Som individuelle egenskaber gælder såvel subkategorisering, dvs. ordets syntaktiske forbindelighed,

som selektionsregler, dvs. ordets semisk/sememiske forbindelighed. Det er velkendt, at syntaktiske egenskaber lettere lader sig systematisere og beskrive i regler end semantiske egenskaber.

Interessant er det i denne sammenhæng at iagttage de bestræbelser inden for den moderne leksikografi, der har til hensigt at afklare, hvilke grammatiske oplysninger der er brug for i store monolingvale ordbøger (jf. fx bidragene i Bergholtz/Mugdan 1985).

### 2.3 Støtteverbumbonstruktioner

Som nævnt ovenfor (i afsnit 2.2) udgør flerordsenheder et stort problem i maskinoversættelse, især når systemet er multilingvalt og ikke kan udnytte "tilfældige" ligheder mellem to bestemte sprog. Men der er en velafgrænset gruppe af flerordsenheder, som det inden for EUROTRA er lykkedes at behandle på en både teoretisk og praktisk tilfredsstillende måde, nemlig en konstruktionstype, som på engelsk kaldes "support verb constructions". Vi fordanser det til "støtteverbumbonstruktioner", men bruger for nemheds skyld også forkortelsen SVC. Det drejer sig om udtryk som:

træffe en beslutning om noget  
foretage en undersøgelse af noget.

Disse konstruktioner henregnes i reglen til kollokationer (altså leksikalske), men kan beskrives med et system, der ligner valensbeskrivelsen, og ligesom den egentlige valens er systemet implementeret i EUROTRA og har vist sig at fungere.

Når man inden for EUROTRA har valgt først at koncentrere sig om denne konstruktionstype, skyldes det to ting. For det første er konstruktionstypen meget almindelig i alle EF-sprogene, specielt i administrative tekster, og giver ofte problemer ved oversættelse. For det andet har det vist sig muligt at lave definitioner og beskrivelsesmodeller, som er anvendelige for alle sprogene.

SVC'erne er ikke identiske med det, der i den germanistiske tradition kaldes "Funktionsverbgefüge", men er groft sagt en delmængde af dem. Kernen i en SVC er et substantiv, som har en argumentstruktur, altså er valensbærende. Typisk vil der være tale om et logisk subjekt (kaldet argument 1) og et logisk objekt (argument 2), fx:

Folketingets (arg1) *beslutning* om nye skatter (arg2)

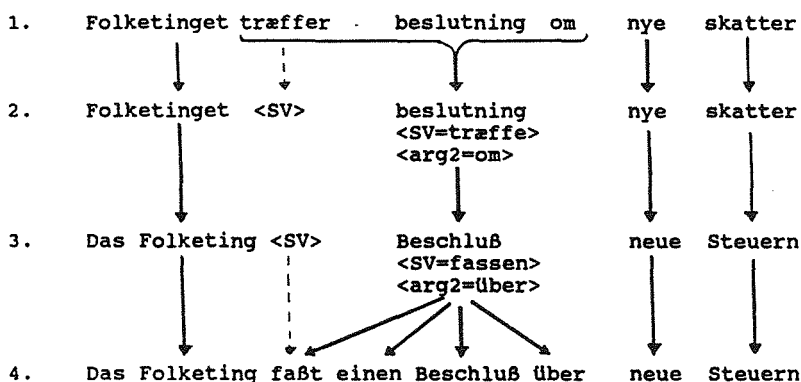
Hvis et sådant valensbærende substantiv står som objekt for et verbum, som i konstruktionen er uden egentligt leksikalsk indhold, kaldes verbet for et *støtteverbum* og helheden for en SVC, fx:

Folketinget *træffer* beslutning om nye skatter

Oplysningen om, hvilket støtteverbum et substantiv forbindes med, anbringes i ordbogsindgangen for substantivet. Det gør, at man i oversættelsesordbøgerne slipper for at oversætte støtteverberne, hvilket ellers er en meget besværlig opgave ved maskinoversættelse, fordi målsprogsverbet ikke er forudsigeligt ud fra kildesprogsverbet, men bestemmes af substantivet på målsproget.

Når substantivet er oversat, kan man derfor blot hente det rigtige støtteverbum i ordbogsindgangen for substantivet på målsproget, hvor man også henter oplysninger om bl.a. valensbundne præpositioner.

Stærkt forenklet ser oversættelsesprocessen ved oversættelse fra dansk til tysk altså ud som vist i fig. 2:



Figur 2

Der er mange aspekter i forbindelse med SVC'er, som vi ikke kan komme ind på her, fx hvorledes man har løst problemet med oversættelse mellem SVC'er og simplex-verber, men et væsentligt formål med vores projekt er at gøre det muligt at beskrive andre typer flerordsenheder tilsvarende eksplicit og implementerbart.

### 3. Metode

Der findes på dansk et opslagsværk, som - uden at begrænse sig til flerordsenheder - indeholder en ret omfattende samling af den slags udtryk, nemlig Erik Bruuns *Dansk Sprogbrug* (DS). Vi har derfor fundet det naturligt at tage udgangspunkt i denne bog. Ud over - som nævnt i afsnit 1 - at forsøge at underbygge det teoretiske fundament for typologisering og klassifikation af flerordsenheder består arbejdet af følgende faser:

- 1 En undersøgelse af, hvordan DS beskriver flerordsenheder: Hvordan er makro- og mikrostrukturen? Hvordan er variabiliteten angivet, først og fremmest m.h.t. syntaktiske og semantiske oplysninger om de variable elementer?
- 2 En empirisk undersøgelse af, hvilke eksempler på flerordsenheder der findes i et større maskinlæsbart dansk korpus (DK87-90)<sup>11</sup>.
- 3 En sammenligning af korpuseksemplerne med beskrivelsen i DS og i et vist omfang andre opslagsværker med henblik på at skabe det empiriske grundlag for næste fase.

- 4 En udtømmende, eksplicit og entydig registrering af disse enheders syntaktiske og semantiske egenskaber, specielt deres variabilitet. Registreringen udformes således, at den kan bruges i maskinel analyse og oversættelse, men samtidig kan omsættes til en "menneskevenlig" beskrivelse i en aktiv ordbog for mennesker.

Faseopdelingen skal ikke forstås således, at fase 1 afsluttes, før fase 2 påbegyndes osv., men således, at faserne gennemløbes flere gange. Vi har afsluttet første gennemløb, der nærmest har haft karakter af et pilotprojekt, idet vi har gennemløbet faserne med et begrænset materiale, og bruger i øjeblikket bl.a. resultaterne af dette arbejde til at justere det teoretiske apparat.

Vi har altså valgt at tage udgangspunkt i DS's systematisering, der er baseret på opslagsordets syntaktiske funktion i den pågældende flerordsenhed. Denne typologi er velegnet som grundlag for vore undersøgelser, fordi den er så overskuelig, at den også kan appliceres på store datamængder såsom ekscerpter fra *DK87-90*. På den anden side er den så grovmasket, at der i det videre arbejde må foretages yderligere opdelinger.

Desværre indeholder de indledende afsnit i DS ingen oplysning om, hvilke overvejelser der ligger til grund for udvalget af de sprogbrugseksempler, der indgår i denne "stil- og konstruktionsordbog" (som undertitlen lyder), altså om forfatteren har haft en anden definition (eller intuitiv opfattelse) af, hvad der udgør flerordsenheder, eller om han bevidst ikke har ønsket at skelne mellem flerordsenheder og frie forbindelser. I hvert fald er det tydeligt, at DS anfører eksempler på begge dele uden at markere nogen skelnen mellem dem. At få noget nærmere at vide om indsamlingsmetoden og -kriterierne ville også have været af interesse for vort projekt. Vi ved, at der ligger en del yderligere materiale hos forfatteren i mere eller mindre bearbejdet form. Frasorteringen af dette materiale skete af redaktionelle hensyn.

Systematikken i DS har to sider: For det første omfatter den ordbogsartiklernes struktur: "Eksempler på ensartede anvendelser er samlet under samme bogstavbetegnelse", dvs. at ordbogsartiklen er delt op i afsnit på grundlag af opslagsordets syntaktiske funktion:

Ved substantiver og adjektiver:

- a: som subjekt (formelt og egentligt);
- b: som objekt (direkte og indirekte) samt prædikativ;
- c: som styrelse for præposition og efter konjunktion;
- d: som underled;
- e: med underled og ved sideordning.

Ved verber:

- a: med subjekt;
- b: med objekt og prædikativ;
- c: med adverbial;
- d: som underled;
- e: som styrelse, [ikke-verbalt] sætningsled samt ved sideordning.

For det andet angives flerordsenhedernes variabilitet, altså hvilke led der er obligatoriske og fakultative led; hvilke elementer der er faste og variable; samt de enkelte leds alternationsmuligheder. Dette sker ved hjælp af et markeringssystem med parenteser og skråstreger:

- [ ] for fakultativ komponent;
- [ ] [ ] sidestilling af to eller flere fakultative komponenter;
- ( ) obligatorisk komponent, eksemplificeret;
- [( )] fakultativ komponent, eksemplificeret;
- / alternerende komponenter (enkeltord);
- // alternerende komponenter (ordforbindelser);
- skilletegn inden for samme afsnit, der kan omfatte flere syntaktiske funktionstyper.

### 3.1 Bearbejdning af DS-materiale

Vi har udvalgt nogle artikler fra DS, der umiddelbart så ud til at rumme mange forskellige slags ordforbindelser. Disse artikler (med opslagsordene *hånd, hoved, slutte, holde, hjælp, hjælpe* m.fl.) har vi fået adgang til i maskinlæsbar form.

Artiklerne er blevet foldet ud, dvs. at den komprimering, der er nødvendig af pladshensyn i en trykt ordbog, er blevet ophævet: Tilde og tankestreg er blevet erstattet med opslagsordet, og i alle tilfælde, hvor der er angivet alternative eller fakultative udfyldninger af en plads i et udtryk, er det samlede udtryk anført én gang for hver udfyldningsmulighed. Denne udfoldning, som kun delvis kan automatiseres, er nødvendig for at muliggøre en følgende maskinel strengsammenligning med korpusbelæg.

### 3.2 Korpusarbejde

Som nævnt ovenfor, er det korpus, vi har anvendt, DK87-90, som rummer 1 million løbende ord pr. årgang, altså i alt 4 millioner ord<sup>12</sup>. Ved hjælp af et søgeprogrammel excerperede vi alle forekomster af et givet ord i alle bøjningsformer. Eksemplerne blev hentet som KWIC (Key Word in Context) belæg, med tilstrækkelig lang venstre og højre kontekst, således at belæggene altid gav meningsfyldte eksempler på søgeordets brug.

Det viste sig, at vi fik en temmelig stor mængde data ud for hver søgning (i alt ca. 3 MB). Bearbejdelsen af disse data var ganske tidkrævende. Vi vil her kun komme ind på eksemplerne med ordet *hånd*.

Vi har undersøgt ordet og dets kontekst systematisk ud fra følgende overordnede kriterier:

- \* Sortering af eksempler i følgende grupper:
  - frie ordforbindelser
  - okkasionelle konstruktioner, herunder også kontaminationer
  - ægte flerordsenheder.



- \* Grovtypologi inden for flerordsenheder på grundlag af:
  - opslagsordets (mulige) syntaktiske rolle(r)
  - opslagsordets kontekstpartnere.
- \* Sortering inden for grovtypologiens klasser i to typer:
  - lukkede enheder (med høj leksikalsk stabilitet)
  - åbne enheder (med leksikalsk variabilitet af forskellig art og grad).
- \* Registrering af morfosyntaktiske varianter.

At denne undersøgelse kun kan give et fingerpeg om, hvor mange forskellige kombinationer af variabilitet og restriktioner der findes i flerordsenheder, er vi godt klar over. Det viste sig imidlertid ret tydeligt ved sammenligningen af de bearbejdede korpusbelæg og DS's typologi hhv. markeringssystem, på hvilke punkter der først og fremmest er brug for en videregående analyse. Dette vil blive omtalt ved sammenligningen af de to datasæt.

Det er indlysende, at en så detaljeret gennemgang af det samlede ordforråd vil kræve en meget stor arbejdsindsats. Den her valgte metode er først og fremmest egnet til at danne sig et systematisk overblik.

### 3.3 Begrundelse for metodevalg

Af ovenstående fremgår det, at undersøgelsen primært er syntaksbaseret. Det er almindeligt kendt, at syntaksen er det felt, hvor man kan bedst skaffe sig et systematisk overblik ved overfladeanalyse.

Syntaksen alene er imidlertid på længere sigt utilstrækkelig som grundlag for en systematisering. En semantisk beskrivelse, der kan redegøre for affiniteten mellem ords betydninger, kræver en dyberegående analyse og et mere detaljeret beskrivelsesapparat. I første omgang nøjes vi med at liste de semantiske mønstre, vi umiddelbart kan iagttage ved gennemgang af materialet, til senere brug ved opstilling af et fuldt udbygget beskrivelses-system.

## 4. Sammenligning af korpusbelæg og repræsentation i DS

### 4.1 Den undersøgte datamængde

Den samlede datamængde fra DK87-90 for søgeordet *hånd* er ca. 600 kilobytes. Eksempelvis har vi analyseret 120 belæg med ordformen *hånden* og 220 med *hænder*.

DS registrerer ca. 115 ordforbindelser for dette opslagsord, med markering af deres leksikalske og morfosyntaktiske variationsmuligheder. Dette svarer til ca. 160 ordforbindelser i alt, når man har foldet artiklen helt ud, dvs. når alle varianter og alternativer er medregnede.

Efter at vi havde sorteret og typologiseret korpuseksemplerne, samt registreret eventuelle morfosyntaktiske varianter af samme flerordsenhed, foretog vi en sammenligning af de to datasæt. Det, vi i første omgang har interesseret os for, er at få overblik over, hvor stor en del af de autentiske eksempler på nutidig sprogbrug der er medtaget i DS, og hvordan registreringen er udformet.

## 4.2 Hvad kan vi konkludere af sammenligningen?

### 4.2.1 Dækningsgraden

Med hensyn til dækningsgraden mellem DS og de undersøgte korpuseksempler har vi kunnet konstatere, at der i DS mangler en del almensproglige flerordsenheder (både kollokationer og fraseoleksemmer), der forekommer flere gange i vort korpus-materiale, fx:

lægge *hånd* på ngt (= beslaglægge ngt.)  
 give *én hånden* på at + inf. (= love én ngt.)  
 tage sagen i egne *hænder* (= handle/bestemme selv)  
 holde *hænderne* fra én/ngt. (fig., = holde sig fra én/ngt.)

Dette kan have flere grunde:

- \* Leksikografens personlige skøn om, hvad brugeren hyppigst vil søge i ordbogen, dvs. om brugseksemplet var nødvendigt at medtage i ordbogen eller ej. Da ordbogens maksimale omfang på forhånd var fastlagt af forlaget, kunne det indsamlede materiale ikke medtages i sin helhed.
- \* De pågældende udtryk er af leksikografen ikke betragtet som leksikaliserede eller på anden måde specielle ordkombinationer, men som frie ordforbindelser. Dette spørgsmål kan vi ikke komme videre med uden at kende leksikografens klassifikationskriterier.
- \* Tilfældigt 'hul' i DS's belægsamling.
- \* Udtrykket er en nydannelse, der først er opstået eller leksikaliseret, efter at materialeindsamlingen er afsluttet.

Det skal tilføjes, at der selvfølgelig også er en del af de flerordsenheder, der er registreret i DS, som ikke er belagt i vores korpus, men vi har ikke anset det for relevant at lave en egentlig statistisk sammenligning.

### 4.2.2 Registrering af leksikalske varianter i DS

I denne henseende er forholdet mellem de to datasæt nogenlunde det samme som under det foregående punkt.

Som nævnt i afsnit 3.2 inddeler vi flerordsenhederne i to hovedgrupper med hensyn til deres leksikalske variabilitet:

Type a	Lukkede enheder (Det variable element har kun ganske få alternativer)
Type b	Åbne enheder (Det variable element har en række alternativer, eksempelvis over 3 eller 4. Tallet er i princippet ikke afgørende; man kan justere det efter bearbejdelsen af en større datamængde, hvis man konstaterer, at et andet tal er mere hensigtsmæssigt.)

DS oplyser ikke noget om, hvor mange og hvilke udfyldninger der er mulige for obligatoriske led, som er angivet i rund parentes, fx

DS: (*gå*) fra hånd til hånd.

DS: (*børnene/sagen*) er i gode hænder

I det første eksempel viser korpusbelæggene, at listen bør udvides til: '*gå/vandre/blive rakt*', men efter vores skøn er der næppe flere udfyldningsmuligheder. Udtrykket hører derfor til de lukkede kollokationer (type a) og burde have været anført som alternerende komponenter (uden parentes).

I det andet eksempel kan eksempelvis også (... *dokumenterne / ejendommen / bedstemor/...*) være alternativer. Udtrykket er altså en åben kollokation (type b).

DS dækker ikke - og kan af pladshensyn heller ikke dække alle individuelle kombinatoriske egenskaber for hvert opslagsord. Men vi har konstateret en vis mangel på konsekvens i den anvendte præsentationsmetode. Det kan illustreres med et enkelt eksempel fra DS, der indeholder begge ovenfor nævnte typer af leksikalsk variabilitet.

DS: *have/holde* (en *stok/en økse*) i hånden

Det verbale led er anført med to varianter som alternerende komponenter (type a), hvorimod det obligatoriske objekt markeres og præsenteres som en liste med to eksempler (type b).

For den første hovedgruppe (type a) gælder det, at det er muligt og ønskeligt at præsentere alle leksikalske alternativer for et udskifteligt element. Udover '*have/holde*' har vi et ret stort antal korpuseksempler med '*tage*' som verballed, men der optræder ikke andre verber på denne plads. Man kan derfor sige, at det verbale leds alternativer danner en endelig (og meget begrænset) liste, men at et af alternativerne mangler i DS.

For den anden hovedgruppe (type b) gælder det, at det udskiftelige element har et utal af alternativer. I den ovenfor viste flerordsenhed danner objektets mulige udfyldninger en åben liste, der kan omfatte meget forskellige substantiver, hvilket også er tilfældet i vore korpusbelæg. Alternativerne har dog visse fælles overordnede semantiske træk, fx '*genstand af en for et menneske håndterlig størrelse*'.

Konklusionen må være, at skellet mellem disse to listetyper bør opretholdes i præsentationen i ordbogen. Selv om ordbogens markeringssystem tillader en vis skelnen og systematik, bliver principperne ikke altid overholdt og kan dermed ikke tages for fulde pålydende.

Begge slags lister bør så vidt muligt baseres på et bredt sammensat stort korpus, dels for at undgå huller i materialet, dels for at udelukke, at okkasionelle ordforbindelser bliver håndteret som standardsproglige.

Den lukkede liste affattes eksplicit og udtømmende på grundlag af belæggene. Den åbne liste kan ikke være udtømmende på samme måde, selv ikke med basis i meget store datamængder. En sådan liste baseres derfor på en beskrivelse af alternativernes semantiske egenskaber, dvs. listens elementer bliver ikke leksikalske individer, men semantiske typer. Listen konstrueres ved at udtrække alternativerne fra korpusbelæggene for det variable led og undersøge, hvilke fællestræk der er karakteristiske.

### 4.2.3 Morfologisk og syntaktisk variation

DS markerer generelt de vigtigste variationsmuligheder, fx om substantivet kan anvendes i både ental/flertal, ubestemt/bestemt form, eller om der kan optræde alternerende præpositioner på samme plads.

Ved sammenligningen mellem de to datasæt viste det sig, at registreringen heller ikke på dette punkt er helt systematisk og entydig og heller ikke udtømmende. Det gælder fx numerusbrugen:

DS: stikke *hånden/hænderne* i lommen vs.  
slå ud med *hånden*.

I vore belæg er formen '*hænderne*' hyppigere end entalsformen i den sidstnævnte konstruktion, følgelig burde denne alternering også være anført.

Det er også noget uigennemskueligt, hvilke kriterier der er afgørende for, om DS bringer det verbale led i en flerordsenhed i grundform (infinitiv) eller i bøjet form (præsens eller præteritum), altså hvad der fx er grunden til den forskellige behandling af:

DS: (børnene *går*) hånd i hånd vs.  
(*gå*) fra hånd til hånd, eller  
(*stå*) med hænderne i siden.

I nogle tilfælde er valget af en speciel finit form velbegrunderet. Det er således naturligt at bruge perfektum i det følgende eksempel, selv om andre former er mulige:

DS: (brevet) *er kommet* mig i hænde.

Problemet er bl.a., at der ikke er foretaget en systematisk skelnen mellem, om de morfologiske og/eller syntaktiske restriktioner for en given flerordsenhed gør en "normalisering" til en (sub)lemmaform urimelig, eller om flerordsenheden anføres som et aktualiseret brugs-eksempel. Denne skavank, som DS har til fælles med mange andre ordbøger<sup>13</sup>, er fatal for NLP og er med til at gøre livet besværligt for den, der oversætter til et fremmedsprog.

Konklusionen for dette punkt er, at registreringen af morfologisk og syntaktisk variation bør være ensartet og udtømmende. Dermed menes, at variation altid skal markeres, hvis den er restringeret eller opfattes som afvigende i forhold til de generelt gældende grammatiske regler. Hvis der ikke er restriktioner for et ledes variabilitet i forhold til frie ordforbindelser af tilsvarende struktur, skal leddet registreres i grundform (verber i infinitiv, substantiver i ental/nominativ osv).

## 5. Projektets næste fase

Det er planen nu at foretage en sammenligning af et større datamateriale fra de to datasæt, for at sikre os et solidt arbejdsgrundlag for den afsluttende fase, hvor arbejdsindsatsen vil blive koncentreret om at udarbejde selve beskrivelsessystemet.

Projektet vil også kunne give en vurdering af, hvorvidt materialet i DS vil kunne bearbejdes og udbygges til at kunne opfylde kravet om en moderne "stil- og konstruktionsordbog".

Perspektiverne i korpusbaseret ordbogsarbejde er lovende også til dette formål: de enkelte ords muligheder for at indgå i større meningsenheder eller brugsmæssige enheder kan bedst kortlægges ved undersøgelse af en omfattende mængde autentiske brugseksempler.

## 6. Afsluttende bemærkninger

Med visse forbehold kan det siges, at DS er en god specialordbog: den indeholder det mest omfattende, for mennesker systematiserede materiale om moderne dansk sprogbrug. Når vi i vor undersøgelse er stødt på en del tilfælde af inkonsekvent registrering og præsentation, skyldes det givetvis flere faktorer, som for eksempel:

- \* Ordbogen er lavet på grundlag af en personlig seddelsamling; der kan have været tilfældige huller i dette materiale;
- \* Af pladshensyn var leksikografen nødt til at frasortere en hel del materiale, dette bevirker også, at der mangler visse almindelige flerordsenheder;
- \* Den dengang anvendte, traditionelle leksikografi har væsentlige kvantitative begrænsninger i forhold til de datamatiske metoder, leksikograferne har til deres rådighed i dag, hvilket også har påvirket det kvalitative aspekt.

Vi har i vor undersøgelse haft adgang til relativt store datamængder og avancerede datamatiske hjælpemidler. Desuden stiller vi i dag - sammen med mange andre - anderledes krav til ordbøger, end det var muligt for 30 - 40 år siden.

Sidst men ikke mindst er vore synspunkter præget af, at vi til daglig beskæftiger os med datamatisk behandling af naturlige sprog (NLP). Dette arbejde gør os bevidst om, at der er stort behov for, at sproget beskrives systematisk, eksplicit og entydigt, men også om, at dette ideelle mål bliver meget svært at nå.

Konsistens og konsekvens er nøgleordene, der også bliver fremtidens krav i almen leksikografi, blandt andet fordi materialet i moderne ordbøger skal kunne udnyttes til mange forskellige formål.

## Noter

1. Der skal ikke her gås i detaljer med, hvilke resultater opslagene gav. Der er tale om en række af de svagheder, som er påvist i Braasch (1987) og (1988) og Kjær (1986) og (1987): bl.a., at bøgerne giver forskellige, til dels modstridende oplysninger, og at det ofte er uklart, hvor meget der hører med til det faste udtryk, og hvor meget der er en frit udskiftelig del af et brugseksempel.

2. Som typisk eksempel kan citeres fra H. Mattutat, *Deutsche Grundsprache* (1969) under opslagsordet *der Schuh*: [...] Wdg: jdm die Schuld in die Schuhe schieben, fig (geben) [...]

3. Det ville nærmest betyde "at beskyldte nogen for skylden". (Ud over et "historisk" eksempel i Ordbog over det danske Sprog har vi ganske vist et belæg med "skylden" fra Berl. Tid. fra 20.3.91, men det er karakteristisk nok i en artikel fra bladets Bonn-korrespondent, som efter mange år i Tyskland jævnlig bruger germanismer).

4. DS beskriver altså den gængse brug, men er for kategorisk med formuleringen (under *sko sb*): c. tm. [...]; skyde én i -ene at + sætn. (: beskyldte én for ngt forkert, navnlig meninger, lave motiver)
5. Fx giver Hausmann i sin artikel om "Collocations" i WDD (Hausmann 1989,1010) blot en ekstensionel definition af begrebet uden at begrunde sin afgrænsning.
6. Se fx Skyum-Nielsen (1991).
7. Se Coseriu (1971) og v. Polenz (1973).
8. Et eksempel på en sådan kreativ leg med ordene i vores førnævnte faste udtryk er overskriften "Was sich Große in die Schuhe schieben", refererende til gaver på St. Niklausdag (Citeret fra Wotjak 1989,461). Det ville næppe være muligt - og i hvert fald ikke rimeligt - at udforme en leksikografisk beskrivelse sådan, at den også tager højde for den slags variation.
9. Se Poulsen (ej publ.) og Skyum-Nielsen (1991). Se også Sonja Poulsens artikel i dette bind.
10. Kildesprogets grænseflade er slutresultatet af kildesprogsanalysen og dermed input til den egentlige oversættelse (transfer). Målsprogets grænseflade er outputtet fra transfer og dermed input til genereringen af overfladeteksten på målsproget.
11. Se Bergenholtz (1988) og DK87-90 i litteraturlisten.
12. Se note 11 ovenfor.
13. Se her igen de førnævnte arbejder af Braasch og Kjær.

## Litteratur

- BBI = Benson, M. / E. Benson / R. Ilson. 1986. *BBI Combinatory Dictionary of English. A Guide to Word Combinations*. Amsterdam - Philadelphia
- WDD = Hausmann, F.J. / O. Reichmann / H.E. Wiegand / L. Zgusta (Hrsg.). 1989ff. *Wörterbücher - Dictionaries - Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Bd. 1-3. de Gruyter, Berlin - New York
- Bergenholtz, Henning. 1988. DK87: Et korpus med dansk almensprog. I: Henning Bergenholtz / Finn Frandsen (red.) Hermes, Tidsskrift for Sprogforskning nr. 1 - 1988 s. 229-237. Det erhvervsproglige Fakultet, Handelshøjskolen i Århus
- Bergenholtz, Henning. 1989. *Probleme der Selektion im allgemeinen einsprachigen Wörterbuch*. Kap. 68 i WDD Bd. 1 (S. 772-779).
- Bergenholtz, Henning / Joachim Mugdan (Hrsg.). 1985. *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28.-30.6.1984*. (Lexicographica. Series Maior Bd. 3). Tübingen
- Bork, Egon. 1990a. *Dansk-tysk Ordbog*. (Gyldendals Røde Ordbøger) 9. udg., 4. opl. Gyldendal, København
- Bork, Egon. 1990b. *Tysk-dansk Ordbog*. (Gyldendals Røde Ordbøger) 12. udg., 4. opl. Gyldendal, København
- Braasch, Anna. 1987. *Zur lexikographischen Kodifizierung von Phrasemen in einsprachigen deutschen Wörterbüchern*. (Unpubl. specialeafhandl., Kbh. Univ.). København
- Braasch, Anna. 1988. *Zur lexikographischen Kodifizierung von Phrasemen in einsprachigen deutschen Wörterbüchern - aus der Sicht eines ausländischen Wörterbuchbenutzers*.

- I: Karl Hyldgaard-Jensen & Arne Zettersten (ed.): *Symposium on Lexicography IV (Lexicographica, Series Maior Bd. 26)*. Niemeyer, Tübingen
- Coseriu, Eugenio. 1971. System, Norm und 'Rede'. I: E. Coseriu: *Sprache, Strukturen und Funktionen*. XII Aufsätze zur allgemeinen und romanischen Sprachwissenschaft. Hrsg. v. Uwe Petersen ... (= Tübinger Beiträge zur Linguistik 2) 2. verbesserte Auflage. Tübingen
- DK87-90 = maskinlæsbart dansk korpus, udg. af H. Bergenholtz, Århus. (Kan rekvireres ved henvendelse til H. Bergenholtz). Se Bergenholtz (1988)
- DS = Bruun, Erik. 1978. *Dansk Sprogbrug. En stil- og konstruktionsordbog*. (Gyldendals Røde Ordbøger) 1. udg. Gyldendal, København
- Fleischer, Wolfgang. 1982. *Phraseologie der deutschen Gegenwartssprache*. 1. Aufl. Bibliographisches Institut, Leipzig
- Hausmann, Franz Josef. 1989. Le dictionnaire de collocations. I: *WDD* Bd. 1 S. 1010-1019.
- Heid, Ulrich / Sybille Raab. 1988. Kollokationen im Wörterbuch für multilinguale Generierung und maschinelle Übersetzung. I: *LDV-Forum* Bd. 5 Nr. 4 -Sept. 1988 S. 72-73 (Sammendrag af artikel tilsendt medl. af Arbeitskreis Lexikographie der GLDV)
- Kjær, Anne Lise. 1986. *Fraseoleksemmer i tyske étssprogsordbøger*. ARK 29 (Sprog-instituternes Arbejdsrapporter, Handelshøjskolen i København - bearbejdet cand.ling.merc.-specialeafhandling). København
- Kjær, Anne Lise. 1987. Zur Darbietung von Phraseologismen in einsprachigen Wörterbüchern des Deutschen aus der Sicht ausländischer Textproduzenten. I: Jarmo Korhonen (Hrsg.): *Beiträge zur allgemeinen und germanistischen Phraseologieforschung* S. 165-181. Oulu
- Kjær, Anne Lise. 1990a. *Normbetingede ordforbindelser i tysk juridisk sprog*. ARK 56 (Sproginstituternes Arbejdsrapporter, Handelshøjskolen i København - Ph.D.-afhandling). København
- Kjær, Anne Lise. 1990b. Phraseology Research - State-of-the-Art. Methods of describing word combinations in language for specific purposes. I: *Journal of the International Institute for Terminology Research IITF*, Vol. 1 (1990), no. 1-2. Wien
- Kromann, Hans-Peder. 1989. Zur funktionalen Beschreibung von Kollokationen und Phraseologismen in Übersetzungswörterbüchern. I: Gertrud Gréciano (Ed.): *Europhras* 88. *Phraséologie Contrastive*. s. 265-272. Strasbourg
- Pfeffer, J. Alan. 1979. Drei deutsche Korpora im Institut für Grunddeutsch. Ziel und Stand ihrer Auswertung. I: Bergenholtz H. / B. Schaefer (Hrsg.): *Empirische Textwissenschaft* S. 307-323. Scriptor, Königstein/Ts
- Pilz, K.D. 1978. *Phraseologie. Versuch einer interdisziplinären Abgrenzung, Begriffsbestimmung und Systematisierung unter besonderer Berücksichtigung der deutschen Gegenwartssprache*. Göppingen
- v. Polenz. 1973. Sprachkritik und Sprachnormenkritik. I: G. Nickel (Hrsg.): *Angewandte Sprachwissenschaft und Deutschunterricht*. München
- Poulsen, Sonja. 1991. *Word Combinations as Lexical Building Blocks*. (Prisopgave indleveret til Odense Universitet 1991 - endnu ikke publiceret).
- Skyum-Nielsen, Peder. 1991. *Fyndord. Studier i kortformernes retorik*. Bd.1-3.(Udkommer 1991 ell. evt. 1992 hos Hans Reitzel). København
- Svensén, Bo. 1987. *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. (TNC 85) (Især s. 92-105: Kollokationer och idiom). Stockholm

Wotjak, Barbara. 1989. Ansatz eines modular-intergrativen [sic!] Beschreibungsmodells für verbale Phraseolexeme (PL). I: Gertrud Gréciano (Ed.); *Europhras 88. Phraséologie Contrastive* S. 459-467. Strasbourg



Patrick Nigel Chaffey

## Problemer i forbindelse med tospråklig juridisk leksikografi

De viktigste problemer innenfor norsk-engelsk juridisk leksikografi er forbundet med termer som ikke har noen oversettelsesekvivalent på engelsk eller som er delvis ekvivalente med en engelsk term. Ved hjelp av begrepsanalytiske metoder kan man forsøke å komme frem til brukbare løsninger, men på grunn av behovet for standardisering er det viktig at arbeidet med lovoversettelser går hånd i hånd med det leksikografiske arbeidet. Det er påkrevet at et organ som Rådet for teknisk terminologi automatisk får tilsendt ordlister og glossarer, helst i maskinleselig form.

### Innledning

Det er blitt sagt at en tospråklig ordbok egentlig er "a contradiction in terms". Hvis denne påstanden er riktig, kan det neppe være noe annet felt hvor den er riktigere enn nettopp tospråklig juridisk leksikografi. Selv om jeg for tiden arbeider med en norsk-engelsk juridisk ordbok<sup>1</sup>, kan jeg allikevel opptre som *advocatus diaboli* og konstatere at det egentlig er en håpløs oppgave å forsøke å lage tospråklige juridiske ordbøker, ettersom hvert land utvikler sitt eget rettsvesen og sine egne juridiske begrep og termer mer eller mindre uavhengig av utviklingen i andre samfunn, med mindre det er snakk om nært beslektede samfunn fortrinnsvis med nært beslektede språk. En juridisk term er ofte så kulturbundet at det er umulig å finne et noenlunde tilsvarende ord på et annet språk. Dette er et fenomen som Catford (1965) betegnet som "cultural untranslatability". Anderson (1977) tok konsekvensen av dette og laget en spesiell type ordbok, nemlig en sammenlignende encyklopedisk ordbok, hvor han tok for seg sivilrettslige termer på dansk, norsk, svensk og engelsk. Andersons arbeid har vært meget verdifullt, ikke minst fordi han klarla vesentlige forskjeller mellom de termene han behandlet, men for oversettere er hans ordbok dessverre av begrenset verdi, siden den stort sett ikke gir oversettelsesekvivalenter.

Det er ingen tvil om at det er et stort behov for tospråklige juridiske ordbøker, nettopp fordi lovtekster, forskrifter og direktiver må oversettes fra ett språk til et annet eller kanskje flere andre språk. Jeg har for eksempel oversatt den nye utlendingsloven og utlendingsforskriften til engelsk for Justisdepartementet. Det sier seg selv at utlendingsloven er av begrenset interesse for nordmenn, men den er av stor interesse for en god del mennesker som ikke kan norsk. Jeg antar at det er ganske mange som arbeider med juridisk terminologi enten i forbindelse med oversettelse eller med leksikografi. Dette arbeidet foregår innenfor en gitt virkelighetsramme, og jeg tror vi kan fastslå en del kjensgjerninger som i utgangspunktet ikke burde være kontroversielle:

1 Hvert land bruker et språk som er i stand til å tilfredsstille et klart behov.

- 3 Land som England eller Spania har et språk som er polykulturelt, dvs. at språket ikke danner en kulturell enhet i forhold til ett land. Engelsk og spansk er nasjonalspråk i flere land med store innbyrdes kulturelle forskjeller.
- 4 Et majoritetsspråk har vanskelig for å oppta ord fra et minoritetsspråk, dvs. at mens norsk forholdsvis lett opptar engelske ord, har engelsk enorme problemer med å oppta norske ord. Derimot vil engelsk kunne låne ord fra spansk eller fransk, og omvendt.

### Grunnleggende kjensgjerninger

Det som er sagt har klare konsekvenser for vårt arbeid innenfor tospråklig leksikografi. Hvis vi nå trekker konklusjoner med hensyn til fremstillingen av en norsk-engelsk juridisk ordbok, kan vi fremsette følgende påstander:

- 1 Norsk er et språk som tilfredsstiller behovet i Norge. Med dette menes at språket inneholder alle de juridiske termer og uttrykk som er nødvendige for at rettssystemet i Norge kan fungere på norsk.
- 2 Norske juridiske termer lar seg definere i forhold til en norsk virkelighet i Norge. Gulbransen (1987) har laget en encyklopedisk ordbok som inneholder norske definisjoner av termer som betegner en slik norsk virkelighet.
- 3 Engelske juridiske termer lar seg definere i forhold til en virkelighet i et engelsktalende land, men det er flere slike land, og innbyrdes kan disse land være kulturelt svært ulike. Black (1979) definerer juridiske termer stort sett i forhold til USA; Martin (1990) definerer slike termer i forhold til England.
- 4 Det lar seg ikke gjøre å overføre "vanskelige" norske juridiske termer til engelsk. På en eller annen måte er man nødt til å lage en oversettelsesekvivalent. (Her ser jeg bort fra stillingsbetegnelser som *ombudsmann* og *lensmann*, som kan overføres til engelsk, riktignok med engelsk stavemåte.)

De som har størst behov for en norsk-engelsk juridisk ordbok, er de som oversetter norske lovtekster og forskrifter til engelsk. Det er derfor nødvendig å forsøke å lage en ordbok som inneholder oversettelsesekvivalenter, og spørsmålet blir da om det overhodet er mulig å finne tilfredsstillende oversettelsesekvivalenter. På tross av store kulturelle forskjeller mellom Norge og de ulike engelsktalende land finner vi allikevel en del juridiske termer i Norge som lar seg oversette direkte til engelsk uten noen som helst fare for misforståelse, og her er det snakk om *direkte ekvivalens*. Dessverre er direkte ekvivalens heller unntaket enn regelen. Oftest vil det være snakk om *delvis ekvivalens* eller i verste fall om *ingen ekvivalens*. Jeg kan gi noen eksempler på hvert fenomen og kommentere de løsninger vi har valgt.

### Oversettelsesekvivalenter

Når vi har med direkte ekvivalens å gjøre, vil løsningen egentlig være innlysende: man skal simpelthen oppgi den engelske ekvivalenten. Det er klart at det i noen tilfelle kan ta tid å få

brakt på det rene om den engelske termen virkelig er en fullgod ekvivalent, og det kan også ta tid å finne ut om en eventuell engelsk ekvivalent i det hele tatt eksisterer, men når man engang har funnet en slik ekvivalent, så er problemet løst. Jeg kan illustrere dette ved hjelp av følgende termer fra ordboken:

lokalforvaltning	local government
prostituasjon (erhvervsmessig utukt)	prostitution
pågreipe	arrest <v>

Her er det ikke nødvendig å ta med noen definisjon eller forklaring, verken på norsk eller engelsk, fordi man har en norsk term med en klart tilsvarende term på engelsk.

Det motsatte av direkte ekvivalens er ingen ekvivalens, det vil si at man har en norsk term som betegner et fenomen som ikke eksisterer i den engelsktalende verden. Vi kan bruke to norske termer, *rettsvitne* og *sikring*, til å illustrere problemet. For det første er det ikke sikkert at de fleste nordmenn vet nøyaktig hva disse ordene betyr; det er derfor nødvendig å ta med en definisjon på norsk. For det andre er det en norsk virkelighet vi har med å gjøre; det er derfor nødvendig å ta med en engelsk oversettelse av den norske definisjonen.

Når vi vet at en norsk term ikke har noen ekvivalent på engelsk, må vi bruke begrepsanalyse som grunnlag for å danne en ny engelsk term. Det er selvfølgelig ikke uten betenkeligheter man lager en ny term, men dette er tross alt en bedre løsning enn direkte overføring av et norsk ord til engelsk. Sarcevic (1989) har anvendt denne metoden i Jugoslavia, og Newmark (1988) har beskrevet denne metoden under betegnelsen "componential analysis". Gutt (1991) har tatt opp slike problemer ved hjelp av en relevansteoritisk tilnærming til oversettelse. I Norge har Frønsdal og Hansen (1988) brukt en begrepsanalytisk metode under fremstillingen av en norsk-engelsk-tysk spesialordbok. Deres ordbok er et fremragende arbeid som viser at begrepsanalytiske metoder kan anvendes med meget stort hell innen tospråklig leksikografi.

Hvis vi bruker "componential analysis" på termene *rettsvitne* og *sikring*, kan vi konstatere følgende:

rettsvitne	a)	person
	b)	overvårer en rettssak som pådømmes eller behandles av bare én dommer
	c)	skal følge med i forhandlingene
	d)	skal påse at det som føres til protokolls stemmer med det som har foregått
sikring	a)	særreaksjon i strafferetten
	b)	nyttes overfor lovbrøtere med tilbakefallsfare
	c)	regnes ikke som straff
	d)	former oppregnet i straffelovens § 39 nr. 1 a-f

Det er nå vår oppgave å lage engelske termer i samsvar med de ovennevnte komponenter - i den grad det er mulig. Når det gjelder *rettsvitne* er det klart at en slik person ikke er "witness" på engelsk. Vedkommende skal opptre mer eller mindre som en eksaminsinspektør som skal påse at alt går riktig for seg under eksamen. Vi har derfor valgt å lage termen "court invigilator" på engelsk. Med hensyn til *sikring* tenker man automatisk på "preventive detention", men så husker man at sikring ikke regnes som straff. Når man leser straffelovens

§ 39 nr. 1 a-f, blir man klar over at det her er snakk om forskjellige former for tilsyn. Det faller da nokså naturlig å lage den engelske termen "preventive supervision", som faktisk er brukt i den engelske versjonen av *Fangehandboka*.

Vi kommer nå til det kanskje vanskeligste området, nemlig delvis ekvivalens. Her har jeg tenkt å illustrere problemene ved hjelp av de norske termene *voldtekt* og *konkurs*. De fleste oversettere til engelsk vil sikkert spørre med en gang om vi ikke kan fortsette å oversette de to termene med henholdsvis "rape" og "bankruptcy". Det har vi vel gjort i alle år! Problemet er bare det at de norske definisjonene i Norge ikke stemmer med de engelske<sup>2</sup> definisjonene i England og sannsynligvis heller ikke med tilsvarende definisjoner i lovgivningen i de andre engelsktalende land.

Vi kan ta voldtektsbegrepet først. Hva er egentlig voldtekt? Svaret finner vi i straffelovens § 192 første ledd, som lyder:

Den som ved vold eller ved å fremkalle frykt for noens liv eller helse tvinger noen til utuktig omgang eller medvirker til det, straffes for voldtekt med fengsel inntil 10 år, men med fengsel i minst 1 år, dersom den utuktige omgang var samleie.

Det er helt klart at to termer står sentralt i denne bestemmelsen, nemlig "utuktig omgang" og "samleie"<sup>3</sup>. Det er viktig å merke seg at det er den utuktige omgang som er avgjørende for forbrytelsen som kalles voldtekt, ikke det faktum at samleie har funnet sted. Det står heller ikke noe om voldtektsforbryterens eller offerets kjønn. Teoretisk sett kan gjerningsmannen være en kvinne og offeret en mann, selv om voldtekt er en forbrytelse man tradisjonelt regner med begås av menn. Andenæs og Bratholm (1990) sier følgende om begrepet "utuktig omgang":

Utuktig omgang er et begrep som omfatter samleie, men ikke er begrenset til det. Samleie kan ses som en kvalifisert form for utuktig omgang ... Ved siden av samleie omfatter uttrykket utuktig omgang seksuelle handlinger av en viss intensitet. Deltakerne i den utuktige omgang kan være av samme eller av forskjellig kjønn.

Hvis vi når ser på hvordan "rape" defineres i England, finner vi noen vesentlige forskjeller. Mitchell og Richardson (1988) siterer følgende definisjon fra *Sexual Offences Act 1956*:

A man commits rape if

- (a) he has unlawful sexual intercourse with a woman who at the time of the intercourse does not consent to it; and
- (b) at that time he knows that she does not consent to the intercourse or he is reckless as to whether she consents to it

De siterer videre fra *New Zealand Crimes Act 1961*:

Rape is the act of a male person having sexual intercourse with a woman or girl (a) without her consent ...

Black (1979) definerer "rape" på følgende måte:

Unlawful sexual intercourse with a female without her consent. The unlawful carnal knowledge of a woman by a man forcibly and against her will.

Vi ser med en gang at voldtekt i Norge betyr mye mer enn "rape" på engelsk. Hvis en mann voldtar en kvinne i Norge og har fullbyrdet samleie med henne, så er dette "rape" på engelsk; hvis en mann voldtar en kvinne uten å ha samleie med henne, så er det ikke "rape", selv om det er "rape à la norvégienne". Det er det man kaller "indecent assault" i England. Hvis en mann voldtar en mann og har fullbyrdet anal samleie med ham, vil dette tilsvare "forcible buggery" i England. Hvis en mann voldtar en mann uten å ha anal samleie, vil dette igjen falle inn under begrepet "indecent assault".

Problemet er ikke lett å løse. Det er klart at det er en lang tradisjon i Norge for å oversette "voldtekt" med "rape". Vi har derfor nokså motvillig valgt å oversette "voldtekt" med "rape" i ordboken, men vi har tatt med den norske definisjonen og en engelsk oversettelse av denne definisjon. Ved hjelp av "componential analysis" har vi valgt å oversette "utuktig omgang" med "indecent intercourse", selv om uttrykket ikke brukes i de engelsktalende land.

Et lignende problem møter vi når vi skal oversette "konkurs" til engelsk. Tradisjonelt oversettes "konkurs" med "bankruptcy", og Arntzen, Bugge og Underland (1978) har konsekvent brukt "bankruptcy" som oversettelse for "konkurs", som følgende sitat viser:

Persons who have been a member of the board or a managing director of a Norwegian joint-stock company during the last year before the company was declared bankrupt may not ...

Men problemet her er det faktum at ordet "bankrupt" ikke brukes om selskaper i England. Hansen (1988) sier uttrykkelig at man bruker "bankruptcy" om personlig konkurs, dvs. om fysiske personer, mens man bruker "liquidation" eller "winding-up" om selskaper. Her har vi valgt å følge Hansen, og vi har oversatt "konkurs" med både "bankruptcy" og "liquidation". Ved hjelp av definisjoner har vi gjort det klart hva forskjellen er.

Noen vil selvfølgelig innvende at vi har vært inkonsekvente: vi har oversatt "voldtekt" med "rape" på tross av alt som er anført ovenfor, mens vi har valgt å oversette "konkurs" på to forskjellige måter. Ja, vi må bare innrømme at vi har vært inkonsekvente, dog ikke uten grunn. I Norge er voldtektsbegrepet svært omfattende; det dekker flere forskjellige typer straffbare handlinger i England eller USA. Det er derfor praktisk umulig å oversette "voldtekt" med flere forskjellige alternative løsninger, det ville virke forvirrende. Vi er nødt til å si at "rape" i Norge kan, men må ikke, være annerledes enn "rape" i England eller USA. Når det gjelder konkursbegrepet, er det innlysende at dette dekkes av to klare termer på engelsk. Det er vel derfor rimelig å slutte seg til Hansens løsninger.

#### Behovet for standardisering av termer

Jeg nevnte tidligere at det er et stort behov i Norge for oversettelse av lovtekster og forskrifter til engelsk og andre fremmedspråk, og det er meget viktig at en term som forekommer i én lov, må oversettes på samme måte når den forekommer i en annen lov, under forutsetning av at den har samme betydning i begge lovtekster. Ruth Vatvedt Fjeld har påpekt at dette ikke

alltid er tilfelle. Det er derfor nødvendig at oversettelse av lover og forskrifter går så å si hånd i hånd med juridisk leksikografisk arbeide. Det ville være svært uheldig om en tospråklig juridisk ordbok skulle gi andre løsninger enn de løsninger som faktisk er brukt i lovoversettelser, fordi en oversettelse av en lov eller forskrift har i seg selv en kraftig standardiserende virkning. Når det er sagt, må det også sies at det naturligvis finnes lovoversettelser som inneholder uheldige eller mindre heldige løsninger. I 1980 ga Utenriksdepartementet ut en nokså omfattende norsk lovsamling oversatt til engelsk. Disse oversettelsene ble i sin tid samlet av avdøde byråsjef Odd Sandal, som var helt klar over at de er av varierende kvalitet, særlig med hensyn til de terminologiske løsninger som er benyttet. Sandals drøm, som han dessverre aldri fikk tid til å sette ut i livet, var at han sammen med en gruppe oversettere og leksikografer skulle gå gjennom hele samlingen og foreta revisjoner for å få til en enhetlig terminologi.

Det er helt klart at de som oversetter lover og forskrifter, nødvendigvis må kunne ha lett adgang til terminologiske løsninger som allerede er benyttet. Det vil si at et organ som Rådet for teknisk terminologi (RTT) må forsøke å koordinere dette arbeidet, og i den forbindelse vil jeg oppfordre alle som lager ordlister og glossarer til å sende et eksemplar til RTT, helst i maskinleselig form. Kun på denne måten vil vi kunne få til den grad av standardisering som er en tvingende nødvendighet for oversettelse av juridiske dokumenter.

### Eksempler hentet fra databasen på Blindern

Til slutt vil jeg vise et par eksempler fra vår database på Blindern. Som tidligere nevnt blir ordboken fremstilt ved hjelp av databaseprogrammet TPED, laget av Håvard Hjulstad ved RTT. For lesere som ikke kjenner TPED, kan jeg nevne at NB står for bokmål og EN for engelsk. 01 betyr "synonym", 04 betyr "se også" og 31 er en forklaring eller kvasidefinisjon. De norske definisjonene er stort sett hentet fra Gulbransen (1987)

NB00	<u>dokumentfalsk</u>
EN00	1) uttering a forged document 2) forgery
NB31	1) bruk av falsk dokument i rettsstridig hensikt 2) det å forferdige el. anskaffe et falsk dokument el. å forfalske et ekte el. anskaffe et forfalsket dokument for å nytte det el. la det nytte i rettsstridig hensikt
EN31	1) use of a false document for an unlawful purpose 2) making or obtaining a false document or falsifying a genuine document or obtaining a falsified document in order to use it or permit it to be used for an unlawful purpose
NB00	<u>besittelse</u>
NB01	possessio
NB01	innehavelse
NB04	ihendehavelse
NB04	tingsbesittelse
NB04	juridisk besittelse
NB04	rettighetsbesittelse
NB04	kvasibesittelse
EN00	possession

- NB31 etter gjengs teori det faktiske forhold mellom en person og en ting, det faktiske herredømme som foreligger når en person har en viss aktuell rådighet over tingen. Av selve besittelsesforholdet kan det altså ikke sluttet noe om dets materielle berettigelse. Som motstykke til denne definisjon anvendes også uttrykket juridisk besittelse - hvor "besitteren" f.eks. utøver en særlig tinglig rett uten å ha den fysiske rådighet over tingen. Når en bare taler om besittelse menes alltid tingsbesittelse. Istedenfor juridisk besittelse har man også talt om indirekte besittelse (middelbar besittelse) i motsetning til den direkte (umiddelbare eller fysiske) besittelse. Eieren som selv sitter inne med tingen er direkte besitter, men sitter en annen inne med tingen på eierens vegne, er eieren indirekte besitter. Besittelse er et rettslig relevant moment på mange områder, og er bl.a. avgjørende for skillet mellom tyveri og underslag.
- EN31 In theory the factual relationship between a person and an object, the factual power exercised by a person who has a certain existing control of the object. From the fact of possession itself no conclusion can be drawn concerning the material justification for such possession. As a contrast to this definition the term possession in law (juridisk besittelse, rettighetsbesittelse, kvasibesittelse) is also used, where the "possessor" for instance exercises a special right in rem without having physical control over the object. When one simply speaks of possession, actual possession of an object is always meant. Instead of possession in law one also speaks of indirect possession (mediate possession) as opposed to direct possession (immediate or physical possession). If the owner has the object in his possession, he is the direct possessor, but if another person has the object in his possession on the owner's behalf, the owner is the indirect possessor. Possession is a legally relevant factor in many areas, and is for instance conclusive of the distinction between theft and embezzlement.

### Noter

1. Ordboken, som inneholder hovedsakelig strafferettslige termer, skal utgis av Universitetsforlaget i 1992. Den fremstilles på Blindern ved hjelp av databaseprogramet TPED, som er laget av Håvard Hjulstad ved Rådet for teknisk terminologi. *Norsk-engelsk juridisk ordbok* er redigert av Patrick Chaffey og Ronald Walford.
2. Her brukes ordet "engelsk" om England i motsetning til de andre engelsktalende land. Det er viktig å huske på at "English law" betyr engelsk rett eller engelsk lovgivning i England og Wales, ikke Skottland eller de andre engelsktalende land. Jeg har med vilje ikke brukt adjektivet "britisk" i denne artikkelen. Normalt omfatter adjektivet "britisk" hele Storbritannia, dvs. England, Wales, Skottland og Nord-Irland.
3. I 1988 fikk straffeloven en ny § 213 som lyder: "Når bestemmelsene i dette kapittel bruker uttrykket samleie, menes vaginalt og analt samleie."

### Litteratur

- Andenæs, Johs. og Bratholm, Anders. 1990. *Spesiell strafferett*. 2. utg. Oslo.  
 Anderson, R.J.B. 1977. *Anglo-Scandinavian Law Dictionary*. Oslo.

- Arntzen, A., Bugge, J. og Underland, U. 1978. *Company, Trade and Tax Law in Norway*. Oslo.
- (Black, H.C.). 1979. *Black's Law Dictionary*. 5. utg. Minnesota.
- Brækhus, Sjur og Aarbakke, Magnus (red.). 1988. *Norges Lovver 1685-1987*. Oslo.
- Catford, J.C. 1965. *A Linguistic Theory of Translation*. Oxford.
- Frønsdal, Harald og Hansen, Einar. 1988. *Møte- og konferanseterminologi*. Oslo.
- Gulbransen, Egil. 1987. *Juridisk leksikon*. 6. utg. Oslo.
- Gutt, Ernst-August. 1991. *Translation and Relevance*. Oxford.
- Hansen, Einar. 1988. *Bedre engelsk forretningsspråk*. 2. utg. Oslo.
- Juss-Buss. 1987. *Fangehåndboka*. Oslo.
- Martin, Elizabeth (red.). 1990. *A Concise Dictionary of Law*. 2. utg. Oxford.
- Mitchell, Stephen og Richardson, P.J. (red.). 1988. *Archbold: Criminal Pleading, Evidence and Practice*. 43. utg. London
- Newmark, Peter. 1988. *A Textbook of Translation*. London.
- Sarcevic, Susan. 1989. Conceptual Dictionaries for Translation in the Field of Law. I: *International Journal of Lexicography* 4/1989: 277-293. Oxford.
- Utenriksdepartementet. 1980. *Norwegian Laws etc*. Oslo.
- Walker, David. 1980. *The Oxford Companion to Law*. Oxford.



Eli Ellingsve

## Tvetydighet som kommunikasjonsproblem Betydningen av en utvetydig terminologi i internasjonal kommunikasjon

Utvetydig terminologi er av meget stor betydning i nasjonal såvel som i internasjonal fagspråkskommunikasjon. Behovet for ny terminologi øker stadig raskere, i takt med nye tekniske landvinninger og økende internasjonalisering. Og behovet melder seg innenfor stadig flere språkområder. Men også tilsynelatende etablerte internasjonale begrep med terminologi som nærmest oppfattes som en del av allmennspråkets ordforråd, kan føre oss bak lyset i en situasjon der språkbrukere fra ulike land møtes for å utveksle informasjon. *Midtøsten* og *petroleum* kan være eksempel på dette.

I muntlig såvel som i skriftlig kommunikasjon er språk *nøkkelen* til effektiv forståelse. Men språket utgjør også i mange sammenhenger en *barriere* mot forståelse og samarbeid. I dette foredraget vil jeg komme med noen synspunkt på det antatte behovet for

- generell språkkompetanse i profesjonell sammenheng
- særskilte språkkunnskaper for spesielle formål
- standardisering av internasjonale begrep/terminologi

### Generell språkkompetanse i profesjonell sammenheng

Gode språkkunnskaper i faglig sammenheng, kan, paradoksalt nok, fortone seg som mangelfulle kunnskaper i allmennspråksammenheng.

Med god språklig kompetanse i profesjonell sammenheng tenker jeg ikke nødvendigvis på bred kunnskap om et språks struktur og terminologi. Det motsatte er ofte tilfelle: Kvaliteten på fagtekster kan ofte være omvendt proporsjonal med høyt antall synonym, eller forekomsten av varierte, elegante grammatiske konstruksjoner.

Ved kvalitativ vurdering av en allmennspråklig tekst har man tradisjonelt lagt vekt på trekk som "ledig språkføring", "variasjon", "stort ordforråd" osv. Ved vurderingen av en fagspråklig tekst trekker ofte egenskaper som "variasjon" og "stort ordforråd" karakteren ned. I en slik forbindelse snakker man gjerne om "språklig forvirring" og "vaklende terminologi". I svært mange tilfeller etterstrebes en formel- eller frasepreget teksttype, til bruk i håndbøker, bruksanvisninger, lover og regler, prosedyrer osv.

Fenomenet er internasjonalt. En av konklusjonene som kunne trekkes etter en arabisk konferanse om fagspråk, i 1990, understreker dette:

Den arabiske verden ser nå behovet for å ta i bruk ny teknologi, og å bruke språket som et effektivt middel for fagkommunikasjon. Språk kan ikke lenger betraktes som et rent retorisk anliggende, en prydd og en utsmykning, et formål i seg selv i en kommunikasjonssituasjon, styrt av hensynet til språkets musikalitet og dets overflod av synonym (Language International, 1990).

Selv om vi ikke alle har del i en litterær arv som araberne, med ruvende verk som *Koranen* og *Tusen og én natt*, har nok de fleste av oss erfart at de retoriske mønstrene for tekstproduksjon som vi har arvet fra våre nasjonale forfattere, simpelthen ikke kan anvendes når vi blir bedt om å skrive en håndbok, teknisk dokumentasjon, regler eller lover. Fagtekstforfatteren kan ikke uten videre trekke veksler på sine skrivende allmennspråklige kolleger.

### Særskilte språkkunnskaper for spesielle formål

Vitenskapelige og tekniske landevinninger i dette århundret har økt vår viten og vår forståelse av verden rundt oss i et omfang og et tempo som neppe noen kunne ha forutsagt, enn si forestilt seg resultatene av, på forhånd. En økende internasjonalisering er både et resultat av og en drivkraft bak en slik utvikling. Avansert faglig samarbeid er ikke lenger begrenset til hver enkelt bedrift eller foretak, heller ikke isolert bak nasjonale grenser. Samarbeid krever kommunikasjon i enhver forstand, ikke minst verbal kommunikasjon. Behovet for særskilte språkkunnskaper vil også vise seg på flere nivå og på flere fagområder i fremtiden enn hva som er tilfelle i dag.

Med tanke på den raske utviklingen innenfor vitenskap og teknologi, og det universelle preg som disse områdene har, er det neppe for dristig å mene at fagterminologien innenfor disse områdene vil tjene på å bli styrt med en fast hånd dersom ikke det totale kaos skal oppstå. Ved et symposium i German Terminology Society i Tyskland for et par år siden ble det hevdet at fagfolk allerede nå har problem med å diskutere sine egne fag på sitt eget språk. Dette problemet kjenner vi også her hjemme, det er ikke begrenset til andre og større land.

Terminologi spiller en avgjørende rolle når det gjelder tilegnelse og anvendelse av forskningsresultat, f.eks. innenfor tekniske eller naturvitenskapelige fagområder. Fagspråkene i disse områdene krever et stort antall termer for å kunne registrere og beskrive mengden av nyvinninger i alle aspekt.

Utviklingen av ny terminologi for ny teknologi holder dessverre ikke alltid tritt med nyvinningene. Selv om utviklerne eller forskerne lanserer sitt produkt under et bestemt navn eller en viss betegnelse, er det ingen garanti for at denne betegnelsen spres og fester seg sammen med kunnskapen om, eller til og med bruken av, produktet. Det motsatte ser vi ofte eksempel på: vi anvender gjerne ulike "hjemmelagete", ad hoc-pregete private betegnelser på et produkt inntil en lansert betegnelse har vunnet seg et sikrere fotfeste hos tunge brukergrupper.

Men behovet for fagspråkkunnskap er ikke begrenset til kunnskap om ens egen nasjonale terminologi. De siste to årene har f.eks. åpnet et helt nytt potensielt hjemmemarked for europeere, takket være jernteppets fall. De store politiske omveltningene i øst, og den forventete styrkingen av økonomiske alliansebånd i vest i 1992, gjør det nokså sannsynlig at kunnskaper om fremmedspråk, og til dels andre språk enn de tradisjonelle skolespråkene engelsk/tysk/fransk, vil komme til å spille en vesentlig større rolle i Europas fremtid. Også i fagspråkssammenheng vil dette behovet for fremmedspråkkompetanse øke.

Europeerne står overfor et sterkt utvidet hjemmemarked med et problem som tradisjonelt har vært knyttet til fjernere markeder: språkbarrierene. Situasjonen er *her*, og den krever tanker omkring en løsning *nå*.

Ønsket om å beherske et eller flere fremmedspråk i forbindelse med yrkesutøvelse fører ofte til en spesialisert språklæring. Fagspesialister lærer ofte antatt tilstrekkelig fagfraseologi til å kunne diskutere sitt fagområde med likesinnete. De tilegner seg den terminologien som de anser for nødvendig for å kunne kommunisere på et interessant nivå med andre fagfolk. I tillegg utvikler de seg kanskje til språklige roboter, uten evne til å delta i generell samtale på samme interessante nivå.

Her vil jeg reservere meg noe mot en uttalelse fra direktør Riborg M. Ericson ved Norges Eksportskole. I en artikkel med tittelen *Eksportfremmede opplæring - en kritisk investering* (Ingeniørmytt 1991) sies det følgende:

En undersøkelse påviste for få år siden at norske ledesers språkkunnskaper og internasjonale utdannelse og erfaring var klart utilstrekkelige, spesielt tysk og fransk. Når det gjelder engelskkunnskapene, mener Ericson mange nordmenn overvurderer sin kompetanse, som kan være mer egnet for teknisk dokumentasjon og for "small talk" enn for krevende internasjonale forhandlinger.

Jeg tror også at en del nordmenn overvurderer sine fremmedspråkkunnskaper. Men jeg tror at dette også gjelder kunnskapene i fagspråk og evnen til å delta i "small talk" over forfriskningene. Det kan være risikabelt å stille for lave krav i slike sammenhenger, selv om det man sier, ikke nødvendigvis behøver være juridisk holdbart.

### **Standardisering av internasjonale begrep/terminologi**

Under dette punktet vil jeg komme nærmere inn på det som overskriften antyder, nemlig *tvetydighet*.

Det vanlige er gjerne at et leksikalt oppslag betraktes som tvetydig dersom det har mer enn én denotasjon. Jeg vil her utvide grensene for denne definisjonen noe ved å bringe inn et internasjonalt aspekt. Tvetydighet i flerspråklig ordbokssammenheng vil jeg forsøksvis definere slik: *Tilsynelatende ekvivalente nasjonale oppslag i flerspråklige ordbøker er tvetydige dersom de denoterer ulike begrep innen de ulike nasjonale språkområdene.*

I fagspråkssammenheng er det ofte behov for å advare mot følgende "overtro":

- Et begrep er ikke nødvendigvis ett og det samme innenfor to forskjellige språk- eller kulturområder selv om de respektive termer i de to nasjonalspråkene er ordrette oversettelser eller tilnærmet ordrette oversettelser i forhold til hverandre  
(Eksempel: *Midiøsten*)
- Et begrep er ikke nødvendigvis ett og det samme innenfor to språk- eller kulturområder selv om de respektive nasjonale termer er nasjonale varianter av ett og samme lånord (f.eks. fra gresk eller latin).  
(Eksempel: *Petroleum*)

Advarselen kan synes banal og uinteressant. Ikke desto mindre kan kanskje et par eksempel fra geografi og oljeindustri være en nyttig illustrasjon av et problem.

### Midtøsten

Et større norsk konversasjonsleksikon definerer *Midtøsten* slik:

Midtøsten, eng. Middle East, betegnelse som oppstod under den annen verdenskrig som navn på det område som hørte under den britiske militærkommando øst for Suez. Etter hvert som dens område ble utvidet, ble begrepet gitt videre geografisk mening. Noen presis avgrensning av begrepet har man egentlig ikke, men Midtøsten omfatter i vanlig språkbruk

Egypt	Syria
Sudan	Libanon
Libya	Kypros
Israel	Tyrkia
Jordan	Iran
Irak	landene på den arabiske halvøy

(*Aschehoug og Gyldendals Store Norske* 1981)

Tar vi for oss et mindre britisk engelsk oppslagsverk, finner vi en sparsom definisjon av *Middle East* som i hovedtrekk svarer til opplysningene hos A&G:

countries from Egypt to Iran  
(*Oxford Advanced Learner's Dictionary*, 1986)

Så langt så vel. Men så vel er det imidlertid ikke. I følge en anerkjent ordbok for det amerikansk-engelske språk omfatter området *Middle East* følgende:

the area including Iraq, Iran, Afghanistan, and, sometimes, India, Tibet, and Burma

eller

the countries near or east of the eastern Mediterranean, including southwestern Asia (Turkey, Syria, Lebanon, Palestine, Trans-Jordan, Saudi Arabia, etc.)  
(*Webster's New Twentieth-Century Dictionary*, 1983)

Den siste definisjonen får man riktignok under oppslagsordet *Near East*, men med tilvising fra oppslagsordet *Middle East*. Begge termene kan altså åpenbart referere til ett og samme begrep.

Termen - og begrepet - *Nære Østen* er ikke like befestet i norsk allmennspråk som *Midtøsten*. Men en norsk oversetter kan møte *Near East* som en problematisk term, og bli

tvunget til å nærme seg komplekset fra den siden. I følge *Oxford Advanced Learner's Dictionary* (1986) omfatter *Near East*

Turkey etc.

*The Concise Oxford Dictionary* (1964) er en smule mer "concise"/konsis:

Turkey & Balkan States

Dette peker faktisk i samme retning, men vi skal ikke forledes til å tro at vi har grepet om begrepet. Et autoritativt tysk oppslagsverk definerer våre problematiske begrep slik:

*der mittlere Osten*: die südlichen Gebiete Asiens von Iran bis Birma"

*der Nahe Osten*: (= der Vordere Orient) die sozialistischen Länder Osteuropas und die vorderasiatischen Länder

(*Duden Deutsches Universalwörterbuch*, 1983)

Jeg har foreløpig gitt meg her. Det gjenstår ennå å innhente opplysninger om de ulike nasjonale definisjonene av begrepet *Midtøsten* i alle verdensspråk bortsett fra norsk, engelsk og tysk. Jeg er også kjent med at norske historikere, geografer og antropologer vil være interessert i å diskutere Aschehougs og Gyldendals definisjon av begrepet, med referanse til sine fagtradisjoner. Det er ingen tvil om at det her kan finnes mer stoff til ettertanke.

Mens vi venter på den endelige konsensusen, kan vi forestille oss en situasjon der en norsk ingeniør, forretningsmann/-kvinne eller politiker, med generelle fremmedspråkkunnskaper fra norsk videregående skole, ønsker å innhente opplysninger om gassproduksjon, salg, priser og prognoser i Midtøsten fra en informasjonsmedarbeider i et internasjonalt selskap eller pressebyrå. Med sine språkkunnskaper, kanskje gode nok til å kunne produsere en ordrett oversettelse fra norsk til engelsk, eller tysk, og sin nasjonale forståelse av begrepet Midtøsten, hvis term virker så velkjent, kan nordmannen risikere å motta all den informasjon han/hun mangler, men aldri ønsket seg, om markedet i f.eks. Afghanistan, Tibet, Burma, Tyrkia, Albania og Bulgaria.

Kommunikasjonsproblemet er neppe av de helt store. Men det kan føre til unødige tidsbruk og gale tall før det løses. Jeg har sporet en større forståelse i det private næringsliv for økonomiske argument enn for språklige/språkhistoriske når det gjelder behov for språkstøtte og standardisering. Det kunne kanskje være interessant å foreta en kostnadsanalyse i forbindelse med termen *Midtøsten*. Jeg har likevel ingen klare ideer om metode her og nå.

## Petroleum

Et annet eksempel. Ordet *petroleum* stammer opphavlig fra latin; *petra* betyr 'stein', og *oleum* 'olje', altså 'steinolje' direkte oversatt.

En anerkjent amerikansk ordbok for vitenskapelig og teknisk terminologi definerer *petroleum* slik:

A naturally occurring complex liquid hydrocarbon which after distillation yields combustible fuels, petrochemicals, and lubricants; can be gaseous (natural gas), liquid (crude oil, crude petroleum), solid (asphalt, tar, bitumen) or a combination of states.

(*McGraw-Hill Dictionary of Scientific and Technical Terms*, 1989)

Det kan sies en del om denne definisjonen. Blant annet vil det kunne hevdes at ordet *flytende* i definisjonen vil forvirre leseren, siden de *flytende* hydrokarbonene etter det som sies, kan forekomme i alle former; flytende, fast, eller i gassform, eller i en kombinasjon. Forklaringen er muligens at adjektivet *liquid* kan bety noe som er flytende under visse omstendigheter, men ikke alltid. Dette er imidlertid neppe allmennspråkets definisjon av *flytende*.

Ytterligere forvirring kan oppstå i internasjonale fagmiljø dersom den engelske termen *petrol* anvendes i en diskusjon om *petroleum*. Ifølge et annet allmennspråklig autoritativt oppslagsverk kan dette sies om *petrol*:

1. petroleum. (Obsolete)

og

2. gasoline. (British)

(*Webster's New Twentieth-Century Dictionary*, 1983)

For en ikke-innfødt engelsktalende person kan likheten mellom ordene *petroleum* og *petrol* forvirre. Det faktum at *petrol* tidligere også har vært anvendt synonymt med *petroleum*, minsker ikke akkurat språkproblemet (hvor lenge siden/med hvilke konsekvenser/hvilken rekkevidde?).

Til overmål har petroleumsproblemet også en fransk side, som kan bidra ytterligere til den internasjonale forvirringen. På fransk finnes termen *petrole*, som betyr det samme som *petroleum*. Likheten i stavemåte mellom engelsk *petrol* og fransk *petrole* er slående. Igjen er det mulig å forestille seg en situasjon hvor vår nordmann — gjerne med studiebakgrunn fra USA — og en franskmann ønsker å utveksle synspunkt på *petrol* — eller var det *petroleum/petrole/gasoline*? Engelsk er ikke morsmål for noen av dem, derimot det eneste språk de kan bruke seg imellom for fagkommunikasjon. Deres engelske vokabularer kan være helt nødtørftige, og preget av nasjonalt "slag". Samtalen kan følgelig utvikle seg derhen at den som snakker, mottar de svarene som hans tilhører formoder vil være i samsvar med det taleren muligens har ment. Innenfor en norsk kulturkontekst kan en slik samtale best karakteriseres med ordene "Goddag mann, økseskaft".

### Litteratur

*Aschehoug og Gyldendals STORE NORSKE leksikon*. Oslo 1981.

*Duden Deutsches Universalwörterbuch*. Mannheim 1983.

Ingeniørnytt 1991-05-22.

Language International 1990/2.4:14. Mohammed Didaoui, Chief, Arabic Section, United Nations Industrial Development Organization. Vienna.

*Oxford Advanced Learner's Dictionary of Current English*. Oxford 1986.

*McGraw-Hill Dictionary of Scientific and Technical Terms*. 1989.

*Webster's New Twentieth-Century Dictionary*. New York 1983.

*The Concise Oxford Dictionary*. Oxford 1964.

Nina Martola

## Arbetet med Finsk-svensk ordbok: Normeringsproblem visavi finlandssvenskan

En finsk-svensk ordbok som skall kunna användas i Finland kan inte bortse från finlandssvenskan. Att på ett smidigt sätt beskriva två varianter av målspråket är många gånger problematiskt, i all synnerhet som det för den ena varianten saknas en generell enspråkig beskrivning. Artikeln redogör för vilka problem man konfronteras med och hur vi inom finsk-svenska ordboksprojektet har försökt lösa dem.

### I. Inledning

Tvåspråkiga ordböcker omtalas i regel inte i termerna *normativ* och *deskriptiv*. I den mån det görs brukar de väl räknas som deskriptiva. Att klassificera en ordbok som deskriptiv är en fråga om att beskriva *avsikten* med den. Hur den färdiga produkten används ligger utanför redaktionens kontroll.

Det finns automatiskt inbyggt ett visst mått av normativitet i alla ordböcker, också i de rent deskriptiva. Orsaken härtill är att språkbruket till sin natur är normativt.

En modersmålstalare har bestämda föreställningar om hur man får och inte får uttrycka sig i olika situationer. Han har känsla för stilvärden och för regionala och tidsmässiga variationer, och han anpassar hela tiden sitt språk i enlighet med den här normkänslan. Den intuitiva språkbrukaren följer således en norm. Normkänsla är en del av den språkliga kompetensen, både på individplanet och inom hela språkgemenskapen.

En deskriptiv ordbok beskriver språket i bruk, och åtminstone tillsvidare närmast skriftspråket i bruk. Skriftspråket håller sig till förhållandevis fasta regler som inte enbart det intuitiva språkkunnandet styr, utan också i hög grad de skrivregler som skribenterna har fått lära sig. Av de textmassor en ordbok bygger på har största delen dessutom producerats av yrkesskribenter, dvs. journalister, fackboksförfattare och skönlitterära författare, som kan antas ha god kännedom om reglerna.

Det är mot bakgrunden av de gemensamma reglerna och normerna inom en språkgemenskap som orden i en ordbok beskrivs och det är på basis av den som beskrivningen fungerar. Saknas den gemensamma normen blir beskrivningen problematisk. Det är det problemet vi står inför i vissa fall vad gäller finlandssvenskan.

### II. Normer

Ordet *norm* och dess avledningar är mer eller mindre tvetydiga begrepp. I den här artikeln används *norm* ensamt i betydelsen 'deskriptiv norm'. Om en preskriptiv norm avses sägs detta

uttryckligen. *Normativ* och *normativitet* används däremot med avseende på en preskriptiv norm. I rubrikens *normeringsproblem* föreligger bägge betydelseerna.

#### A. Finsk och svensk norm

Finsk-svensk ordbok redigeras utgående från en finsk ordbok i vardande, *Suomen kielen perussanakirja* (=Finsk basordbok). Det är en deskriptiv ordbok med tydliga normativa inslag. Som komplettering används andra finska ordböcker.

Vad gäller den allmänsvenska normen följer vi gängse ordböcker.

#### B. Allmänsvenska — sverigesvenska — rikssvenska

I fortsättningen kommer jag att tala om svenskt standardspråk som *allmänsvenska*, då finlandssvenskan innefattas, eller *rikssvenska*, då finlandssvenskan kontrasteras mot standardspråket. Termen *sverigesvenska* används som sammanfattande benämning på svenskan i Sverige med dess regionala variationer (Liljestrand 1985: 6f, Reuter 1987:3f).

#### C. Den finlandssvenska normen

Det existerar ingenting sådant som en officiellt vedertagen finlandssvensk norm, varken en normativ eller en deskriptiv.

Den allmänsvenska normen är officiell norm också för svenskan i Finland. När det gäller skriftspråket tror jag nog att den väldigt långt kan omfattas av finlandssvenskarna, även om de när de själva skriver omedvetet bryter mot den så pass mycket att man känner igen texter som producerats i Finland. Det finns dock ofrånkomliga skillnader mellan finlandssvenska och rikssvenska, och de skillnaderna måste vi i viss utsträckning beakta.

De finländska särdragen brukar indelas i två typer, *finlandismer* och *fennicism*. Skillnaden är, grovt taget, att fennicismerna skvallrar om finsk påverkan, medan finlandismer är sådant som ålderdomligheter och svenska provinsialismer. Givet är att det inte finns någon skarp skiljelinje mellan typerna.

Det har skrivits ganska mycket om finlandssvenskan; en hel del böcker och ett stort antal allmänt hållna artiklar och en väldig mängd artiklar om enskilda språkdrag. Materialet finns excerperat och registrerat på Forskningscentralen för de inhemska språken, så vi har tillgång till det. Bristen med det är att det nästan genomgående rör sig om kontrastiva undersökningar där finlandssvenskan jämförs med rikssvenskan, och att det mesta kommit till i språkvårdande syfte. En del av de egenheter som presenteras som finlandssvenskheter har i själva verket begränsad utbredning i Finland. Upplysningar om i vilken utsträckning sverigesvenska uttryck eventuellt förekommer jämsides ges heller inte alltid.

Vi saknar alltså en finlandssvensk deskriptiv norm att basera oss på. Det finns inga frekvensundersökningar, inga undersökningar av faktiskt språkbruk i större finlandssvenska textmassor, som skulle kunna åskådliggöra regionala och stilistiska variationer i språkbruket. För våra syften känns materialet om finlandssvenskan därför ofta otillräckligt.



### III. Linjen visavi finlandssvenskan

En och annan frågar sig kanske varför finlandssvenskt språkbruk överhuvudtaget skall beaktas om den officiella svenska normen en gång gäller även svenskan i Finland. Man kan väl nöja sig med de finlandssvenska uttryck som tas upp i SAOL och i övrigt lämna finlandssvenskan därhän?

#### A. Finlandismer som skall med

Om en finsk-svensk ordbok skall kunna användas i Finland måste finlandssvenskan beaktas. Även om man numera kan räkna med en viss marknad också i Sverige, så är det nog ändå i Finland och i finländska förhållanden ordboken kommer att ha sin huvudsakliga användning. Finland är ju ett officiellt tvåspråkigt land med finska som majoritetsspråk, vilket innebär att avsevärda mängden texter översätts.

Det finns ett område där det finlandssvenska bruket *måste* följas och där det rent av finns en officiell norm. De termer som förekommer i lagar och förordningar och som tas med i ordboken måste självfallet överensstämma med dem som används på svenska i Finland. Skiljer sig rikssvenskt och finländskt språkbruk ger vi i finsk-svensk ordbok båda varianterna försedda med vederbörliga etiketter.

Benämningar på en mängd organ, instanser och fenomen i samhället måste också tas med. I en del fall saknas instanserna helt i Sverige, i andra fall finns de men har annat namn.

För det tredje finns det en mängd andra mera informella företeelser som är specifikt finländska och som det givetvis finns behov av att tala om.

Ex.           knackkorv   (små korvar som liknar prinskorv men är längre)  
              boboll        (ett bollspel som påminner om baseboll)

Det vore absurt att föreslå uttryck som "en prinskovsliknande korv" eller "ett baseboll-liknande spel" i finländska sammanhang eller att bara ge närmaste sverigesvenska motsvarighet.

Om det finns mer eller mindre vattentäta skott mellan det rikssvenska och det finlandssvenska uttrycket måste båda tas med. Skulle det stå *korvkaka* i stället för *leverlåda* på en meny i Finland är det är inte många som skulle veta vad det bjuds på.

Orden *knackkorv*, *boboll* och *leverlåda* hör till den grupp finlandismer som kan betecknas som "allmänfinlandssvenska", dvs. det används över hela det finlandssvenska språkområdet och i alla stilarter. Sådana finlandssvenska ord har vi bedömt det som självklart att ta med.

#### B. Gränsdragningsproblem

I vilken mån skall man då ta med andra än allmänfinlandssvenska ord och uttryck, eller med andra ord, vilken tyngd skall man ge finlandssvenskan? Utrymmet är ju begränsat och i praktiken betyder det att varje finlandism finns med på bekostnad av något annat. Risken är att finlandssvenskan får oproportionerligt stort tyngd.

En storordbok innehåller uppslagsord från ett brett stilspektrum och ju ledigare stil desto mera avviker finlandssvenskan från språket i Sverige. Eftersom vår ordbok skall kunna

användas i finländska förhållanden är det nog rimligt att ta med en del utöver de ovannämnda typerna.

Våra användare har behov av att producera texter på olika stelnivåer och för olika målgrupper. Många av de texter som produceras på svenska i Finland riktar sig uteslutande till finländsk publik och är skrivna på ledigt vardagligt språk. Görs en sådan text väldigt rikssvensk kommer den att kännas främmande för det stora flertalet finlandssvenskar och förlorar därmed i informationsvärde.

Det är alltså befogat att bredda urvalet ytterligare. Detta ger dock upphov till många linjedragningsproblem.

## 1. Regionala olikheter

Till att börja med räcker vårt basmaterial, såsom ovan framgått, inte till för ett deskriptivt förhållningssätt. I många fall kan vi inte ens avgöra om ett ord skall anses som allmänfinlandssvenskt. Är det inte allmänfinlandssvenskt betyder det förstås att något annat uttryck används på resten av språkområdet. Detta andra uttryck kan vara antingen en konkurrerande finlandism eller samma uttryck som används i Sverige.

Replikskiftet nedan utspelade sig i en sportaffär i Borgå mellan den infödde expediten och en nyinflyttad österbottnisk kund:

- Jag har kommit för att hämta den där träningsdräkten

- Jaha, var har du den då?

Det till synes obegripliga svaret beror inte på att expediten inte hörde upp utan på en kommunikationsklyfta. Kunden använde *hämta* enligt sin allmänsvenska norm för en rörelse i båda riktningar, medan expediten tolkade det som 'ha med sig' enligt sydfinländsk norm.

Om Borgåbon hade varit kund i en österbottnisk affär kunde dialogen ha låtit ungefär så här:

- Jag har kommit för att söka den där träningsdräkten

- Vad är det för en träningsdräkt? Tror du att den finns här?

*Söka* i betydelsen 'hämta' är sydfinländskt språkbruk och helt främmande för österbottningar.

En ung nylänning skulle förmodligen inte heller ha sagt *träningsdräkt* utan *verrare* (ett lån från finskan).

Man kan grovt taget räkna med tre huvudregioner för finländskt "standardspråk". Det åländska språkbruket är det som ligger närmast sverigesvenskan. Österbotten bildar en region och södra Finland en tredje, båda med avsevärda inre variationer.

## 2. Relationer till rikssvenskan

Förhållandet till sverigesvenskan är invecklat. Vi har självfallet inte alltid att göra med entydiga par av typen *korvkaka* - *leverlåda*. Det förekommer överlappningar och betydelseglidningar av många slag. Det rikssvenska uttrycket kanske förstås men inte används, vilket

är vanligt. Den kan förekomma jämsides på vissa håll, i någon mån överallt eller som en jämbördig konkurrent. Här finns det stora regionala skillnader.

Vi skall se på hur några ord förhåller sig till sin sverigesvenska motsvarighet.

Ordet *villa* avser i Sverige 'enfamiljshus för året-runt-boende'. I Finland används ordet *egnahemshus*, som också har officiell status, dvs. det används i lagar och förordningar. Ordet *villa* är allmänt också bland finlandssvenskarna men uteslutande i betydelsen 'sommarstuga'.

Ordet *batteri* är gemensamt som benämning på batterier i radioapparater, ficklampor, klockor o.likn. I finlandssvenskan har ordet därtill betydelsen 'värmeelement', medan den sverigesvenska betydelsen 'bilbatteri' är främmande. Ett sådant kallas *ackumulator*. Språkvården rekommenderar dock att det rikssvenska spåkbruket skall följas i skrift i detta fall.

Sött vetebröd kallas vanligen *bull* i Finland men ordet *vetebröd* förstås och förekommer också jämsides i vissa regioner.

Finlandssvenskan och rikssvenskan skiljer sig också genom preferens för olika synonymer. Ett sådant par är *våga* och *töras*. Det senare förstås av finlandssvenskarna men används inte aktivt. I Sverige är *töras* det dominerande ordet inom stora regioner och är vanligt i rikssvenskt standardspråk.

De här förhållandena måste beaktas i etiketteringen av ekvivalenter i vår ordbok.

### 3. Vardaglig stilnivå

När man rör sig neråt längs stilskalan ökar skillnaderna, vilket innebär att beskrivningen av den nivån i våra svenska källor ofta gäller uteslutande rikssvenskt språkbruk. En lång rad ord som hör till rikssvenskt vardagsspråk är främmande för finlandssvenskarna. På en del håll förstås de inte ens.

De rikssvenska ord som vunnit insteg i Finland uppfattas ofta som hörande till en annan stilnivå. Orden *kille* och *tjej* t.ex. upplevs sannolikt som slangartade av många finlandssvenskar, medan de för sverigesvenskans del knappast numera kan betecknas som slang. I de här fallen saknas alltså den gemensamma normkänsla som jag talade om i inledningen.

Det är inte bara skillnaderna mellan sverigesvenska och finlandssvenska som blir större när man rör sig neråt på stilskalan. Också de regionala skillnaderna inom finlandssvenskan ökar.

Det här innebär att man för många talspråkliga och slangartade uppslagsord antingen måste nöja sig med de ekvivalenter som finns upptagna i sverigesvenska uppslagsböcker, vilket gör att ordboken i de fallen inte vänder sig till finländsk publik, eller också redovisa flera finländska ekvivalenter med en metaspråklig angivelse om regionen, vilket är att ge en marginell ordtyp stort utrymme. Finlandssvenskan har inte i samma utsträckning som sverigesvenskan ett förråd av vardagliga ord med allmän spridning.

En påfallande regional skillnad är att det på tvåspråkiga områden, särskilt i södra Finland, är vanligt med finska lån och citatlån i ledigt språk. På enspråkiga områden förekommer det mycket sparsamt, och vad mera är, de finska lånen förstås inte ens. Man kan därför inte utan vidare presentera dem som vardaglig finlandssvenska. En intet ont anande översättare som använder dem i syfte att nå ut till den finlandssvenska publiken råkar i stället ut för att inte bli förstörd på vissa håll i landet och för att kanske rent av väcka förargelse.

## C. Etikettering

Vi har försökt utveckla ett system för bruklighetsetiketteringen som beaktar de regionala skillnaderna utan att ge finlandssvenskan opropor­tionerligt utrymme. Vid vattentäta skott mellan rikssvenskt och finlandssvenskt ger vi ett par *sverigesvenska* — *finlandssvenska*. Vanligare är dock att det finns en allmänsvensk ekvivalent, som är gångbar i Finland, och att det vid sidan av den finns en mer eller mindre spridd finlandism. Det allmänsvenska ordet ges först och oetiketterat och vi får alltså ett par  $\emptyset$  — *finlandssvenska*. Den finlandssvenska ekvivalenten kan dessutom få en precisering *finlandssvenska vanligen, finlandssvenska även* beroende på hur spridd den är.

## D. Normativa överväganden

Svenskan är ett minoritetsspråk i Finland (talas som modersmål av knappa 6 %) och existerar under ett enormt tryck från majoritetsspråket. En avsevärd del av de svenska texterna som finlandssvenskarna tar del av är översättningar och, tyvärr, ganska många av dem otypliga och stelbenta. De finländska läsarna kommer att läsa samma fel gång på gång, vilket leder till avtrubbning. Många finlandssvenskar är dessutom tvåspråkiga (bl.a. en så pass viktig påverkargrupp som journalisterna) och lever och verkar i finsk miljö med den risk för uttunning av modersmålet som det innebär. Deras svenska påverkas strukturellt och idiomatiskt av finskan. Följden är att finskan lyser igenom också i texter som produceras direkt på svenska. Finlandssvenskan riskerar att bli svenska på finska.

Vi har bedömt det som vår uppgift att i den mån vi har möjlighet försöka motarbeta de här tendenserna i vår ordbok. I ett land där så mycket översätts kan en storordbok rent av ha ett visst inflytande, åtminstone så pass, att de som tittar efter i ordboken kanske undviker vissa fällor.

Vad gäller de normativa övervägandena arbetar vi i nära samråd med den finlandssvenska språkvården.

Hävdvunna och allmänt spridda finlandismer redovisas. I princip drar vi gränsen vid reglerna för korrekt allmänsvenskt språkbruk. Den felaktiga användningen av *hämta* (ovan) redovisas alltså inte, trots att den är relativt stor spridning.

Finska lånord förhåller vi oss ytterst reserverat till. Vissa mycket användbara och väl etablerade ord får i alla fall komma med (t.ex. *talko* 'arbetsgille'). Direkta lån är inte särskilt vanliga i skrift.

Men det finns en annan typ av fennicismersom är vanliga och dessutom betydligt mera bedrägliga. Det är sammansättningar som är bildade på svenska men enligt modell från finskan.

Fyller de här en nisch kan de accepteras, men ofta är det fråga om onödiga nybildningar som kommit till därför att man inte känt till existerande svenska ord eller inte insett att det betydtt samma som det finska "modellordet". Exempel på sådana ord är *mottagningscentral för flyktingar* eller *flyktingcentral* i stället för *flyktingförläggning* och *hinderritt* i stället för *[ban]hoppning*.

På samma sätt kan hela uttrycksätt lånas in, och trots att uttrycken kan vara nog så målände på ursprungsspråket är det ingen anledning att acceptera dem om de inte förstås av svenskar som inte kan finska. *Att ha sin egen ko i diket* (att ha egna intressen att bevaka, att

tala för sin sjuka mor) och *att lyfta katten på bordet* (att tala klarspråk, att ta tjuren vid hornen, att tala ut om något) är uttryck man då och då ser i finlandssvenska texter.

#### IV. Slutord

Finlandssvenskan och dess förhållande till sverigesvenskan väcker med jämna mellanrum debatt bland finlandssvenskarna. Otvivelaktigt finns det en del påfallande skillnader, och den som följer med språket i bruk får lätt intryck av att skillnaderna håller på att öka. Det är dock lätt hänt att man fäster sig vid skillnaderna och glömmer bort att den överväldigande delen av språkgodset är gemensamt. Faktum är att det är i en ganska liten del av ordboksartiklarna vi behöver ta till etiketten *finlandssvenska*, och då måste man ändå räkna med att skillnaderna mellan sverigesvenskt och finlandssvenskt språkbruk accentueras i en finsk-svensk ordbok, eftersom finskans påverkan på finlandssvenskan aktualiseras just i den. I en engelsk-svensk ordbok anpassad för både Finland och Sverige skulle problemprocenten vara ännu lägre.

#### Litteratur

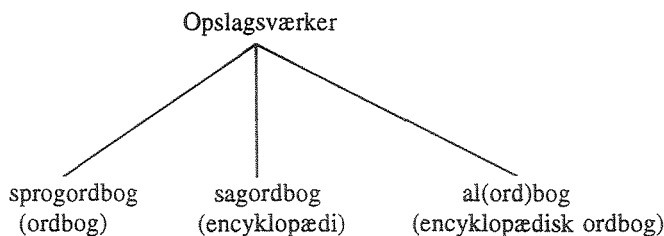
- Liljestrand, Birger. 1985. *Juttu och sammetsfarmare. Notiser om finlandssvenskt tidnings-språk*. Umeå.
- Reuter, Mikael. 1987. Vad är finlandssvenska. I: *Språkbruk 1/1987*: 3-10. Helsingfors.

## Encyklopædiske oplysninger i tosprogede tekniske ordbøger

I forbindelse med oversættelse af tekniske tekster er der et stort behov for encyklopædiske informationer i ordbøgerne. En undersøgelse af den hidtidige praksis på dette område viser, at denne type informationer indtil videre kun gives sporadisk i tosprogede tekniske ordbøger. Der er imidlertid mange muligheder: tekst og tegning, placering i mikro- og makrostruktur samt i ordbogens tekstdel, et system af henvisninger til at knytte informationerne sammen til et hele. En teori for integrering af encyklopædiske oplysninger i tosprogede tekniske ordbøger må udvikles i nær vekselvirkning med den leksikografiske praksis.

Leksikografien deler traditionelt ordbøger (eller opslagsværker) op i to store klasser: sprogordbøger og sagordbøger (encyklopædier). I den første type gives oplysninger om sproget, i den anden om verden (sagen, encyklopædien osv.) Mange ordbøger er imidlertid ikke »rene«, idet sprogordbøger ofte rummer encyklopædiske angivelser, og sagordbøger sproglige angivelser. Hos nogle forfattere betragtes disse »urene« ordbøger som hybrider. Andre klassificerer dem derimod efter deres funktion (»Genuiner Zweck« - Wiegand 1988). Heraf typologiseringen i sprogordbøger, hvis mål er at behandle sproget, men som også kan indeholde oplysninger om sagen, i sagordbøger, hvis mål er at behandle sagen, men som også kan indeholde oplysninger om sproget, og i alordbøger (hos nogle også kaldet encyklopædiske ordbøger), hvis mål er at medtage oplysninger af begge slags. Denne opdeling kan illustreres således:

Skema 1:



Nærværende undersøgelse har ikke til formål at gå ind i denne diskussion, der indtil nu hovedsagelig har kredset omkring ensprogede og almensproglige ordbøger. Den vil derimod anskue problemet i lyset af tosprogede tekniske oversættelsesordbøger, idet den vil fokusere på behovet for at inkludere encyklopædiske oplysninger i sådanne ordbøger, undersøge den hidtidige praksis på dette område og opstille en række principper for det fremtidige arbejde. Undersøgelsen er en del af et 3-årigt projekt, som jeg arbejder med på Handelshøjskolen i

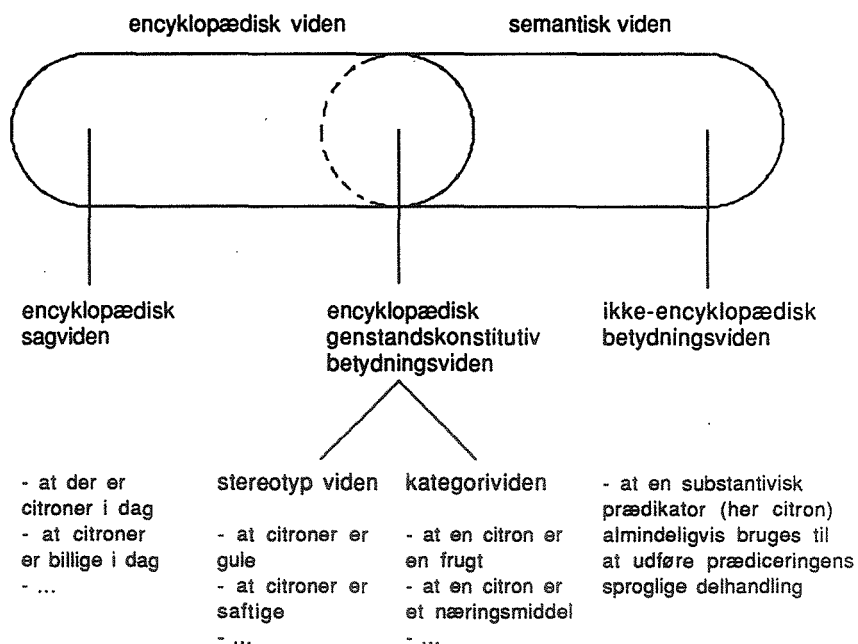
Århus med henblik på at opstille de leksikografiske principper for en dansk-spansk ordbog inden for teknisk sprog.

Da jeg gik i gang med at undersøge dette område, opdagede jeg til min forundring, at det praktisk talt ikke er blevet behandlet i den teoretiske litteratur. Det eneste sted, hvor jeg en passant er stødt på problemstillingen, er hos Scholfield 1980. Derimod er selve diskussionen om forholdet mellem sprog og verden uhyre omfangsrig og synes til tider temmelig ophidset. Synspunkterne placerer sig øjensynlig i et bredt spektrum lige fra den opfattelse, at sproget så at sige kan defineres og forklares ved sig selv, uden henvisning til den ydre verden, til den opfattelse, der får Haiman 1979 til at konkludere, at »dictionaries are encyclopedias«, at der med andre ord ikke findes nogen klar skillelinje mellem sprog og verden.

Wiegand 1988 forsøger at overvinde kløften og har udarbejdet følgende model, idet han ikke mener, at man kan sætte et skarpt skel mellem semantisk og encyklopædisk viden:

Skema 2

Forholdet mellem forskellig slags dagligviden



(Efter Wiegand 1988)

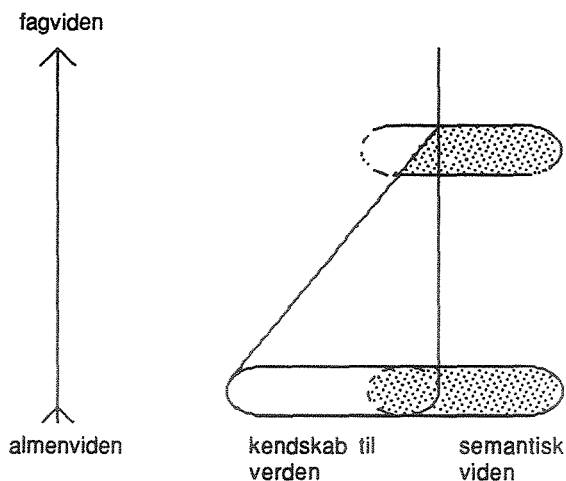
Ser man nærmere på Wiegands model, er det klart, at »citron« her tilhører almensproget (i den snævre betydning: fælles for alle). Forestillede man sig et menneske, som ikke forstod betydningen af ordet »citron«, ville han alligevel ved en kort leksikografisk definition (f.eks.: »en gul, saftig, sur og spiselig citrusfrugt«) kunne slutte bro mellem semantik og encyklopædi, uanset om han havde set en citron tidligere. Forudsætningen for denne enkle operation er lige netop den almenkulturelle viden, som mennesker i et moderne samfund har. Ord som »gul«, »saftig«, »sur«, »spiselig« og i det mindste »frugt« tilhører alle almindelige menneskers dagligsprog.

I sine diskussioner af begrebet fagsprog arbejder Kalverkämpfer ud fra en opfattelse af, at al sprog er fagsprog, blot med en stigende grad af specialisering. Denne stigende specialisering er samtidig for den, der ikke kender de pågældende fag, en stigende grad af uvidenhed. Jo længere han dykker ned i faget, desto færre faste holdepunkter vil han have til at forstå, hvad der tales om.

Disse overvejelser får stor betydning for den leksikografiske definition i fagsproglige ordbøger. Normalt findes der en række muligheder, hvormed man kan give denne definition: parafase, synonym, antonym, hyponym osv. Disse definitioner dækker imidlertid over et langt dybere indhold, som tilsløres i almensproget. Definitionens virkelige kerne er at gengive indholdet i det begreb, som ordet repræsenterer, og placere det i relation til andre begreber, dvs. indordning i et kategorialt system.

Ordets funktion i den menneskelige erkendelsesproces er at generalisere og danne begreber, som kan bruges i kommunikationen og den logiske tænkning, som byggesten for bevidstheden. Straks man ryger op i specialfaget, har lægmand imidlertid ikke kendskab til de konkrete fakta, som ligger til grund for generaliseringen, og han har heller ikke nogen kendskab til eller forståelse for de »nabo«begreber, som definitionen også knyttes til. Problemet kan illustreres i modellen i skema 3.

Skema 3:



Det fremgår her, at den »normale« leksikografiske definition ikke vil kunne løse problemet. Den vil ikke være i stand til at knytte bro ud til verden, idet denne verden ikke er erkendt af



lægmand. Der er således brug for at »hælde« viden på, så definitionen og dens omgivelser kan få form og indhold.

Denne problemstilling gælder ikke blot for ensprogede ordbøger, men også for tosprogede. Man kunne f.eks. i en passiv L2-L1 ordbog betragte oversættelsesordet på modersmålet som en slags »definition«. Det er klart, at en lægmand ikke vil have megen glæde af at få et udenlandsk fagord oversat til et modersmålignende fagord, som han heller ikke forstår. Og bedre bliver det naturligvis ikke i en aktiv L1-L2 ordbog, hvor man tager udgangspunkt i et uforståeligt fagord på modersmålet og får det oversat til et lige så uforståeligt på fremmedsproget.

Man kunne her indvende, at ved en oversættelse er det ikke nødvendigt at forstå teksten. Blot man finder det rigtige oversættelsesord, kan man ved sit grammatiske kendskab til det pågældende sprog foretage en korrekt oversættelse. Praksis viser imidlertid, at sådan forholder det sig ikke. Folk, der arbejder med fagsproglige oversættelser, har for længe siden erkendt, at et minimum af forståelse for teksten (faget) er en forudsætning for en heldig oversættelse. Klarest fremstår problemet i tilfælde af polysemi: hvilken af de opgivne ækvivalenter skal man vælge? Traditionelt er det i fagordbøger blevet løst ved faglige (diatekniske) markeringer, der henviser ordet til det ene eller andet fag eller delområde. Men man kan med rette spørge, hvordan en lægmand skal kunne have glæde af disse angivelser. Det kræver i sig selv faglig viden at placere ordet (teksten) i det rigtige fag (delområde). Det er formodentlig lignende overvejelser, der har affødt følgende bemærkning i forordet til Collazo:

However well grounded in his or her subjects of competence, a translator needs to have access to a reference library including monolingual as well as interlingual dictionaries and encyclopedias, both general and of various degrees of specialization, textbooks and other works in the languages and on the subjects of interest; recent literature on specific systems, devices, techniques, and applications; and other sources of pertinent information and vocabulary.

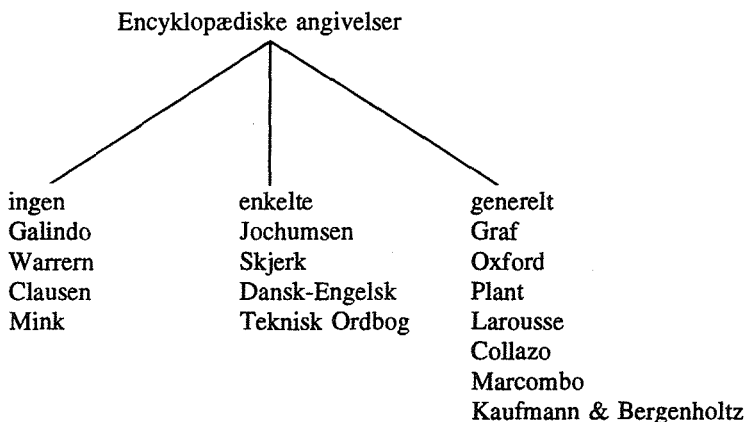
Andetsteds i denne rapport har Maidahl, Laursen og Duvå (s. 105) desuden fremlagt resultaterne af en undersøgelse af problemer, der opstår i forbindelse med fagsproglige oversættelser, og disse resultater peger klart i den retning, at viden om det pågældende fag er en forudsætning for en succesrig oversættelse. Af en ikke publiceret undersøgelse, som Bernt Møller har foretaget på Handelshøjskolen i Århus af problemer ved oversættelse af en informationsteknologisk tekst til fransk, fremgår det endelig, at hele 31% af alle ordbogsopslag drejer sig om forståelse af den faglige sammenhæng i teksten.

Som allerede nævnt har den teoretiske leksikografi praktisk talt ikke beskæftiget sig med spørgsmålet om encyklopædiske oplysninger i fagordbøger. Men hvordan ligger det med fagordbøgerne selv? Hvilken praksis har der udviklet sig på dette felt?

Jeg har foretaget en mindre undersøgelse af den konkrete praksis i 24 en- og tosprogede fagordbøger, som alle har det til fælles, at de står på Handelshøjskolens bibliotek i Århus. Det drejer sig hovedsagelig om tekniske ordbøger, idet jeg dog af forskellige grunde har medtaget en lingvistisk og to juridiske ordbøger. Jeg har i undersøgelsen ikke skelnet mellem ordbøger og encyklopædier ifølge den traditionelle opdeling.

Jeg har først kigget på, hvorvidt der overhovedet er encyklopædiske angivelser med i disse ordbøger. Med encyklopædiske angivelser mener jeg her såvel tegning som tekst, dog ikke faglige markeringer. Resultaterne af undersøgelsen kan illustreres som i skema 4.

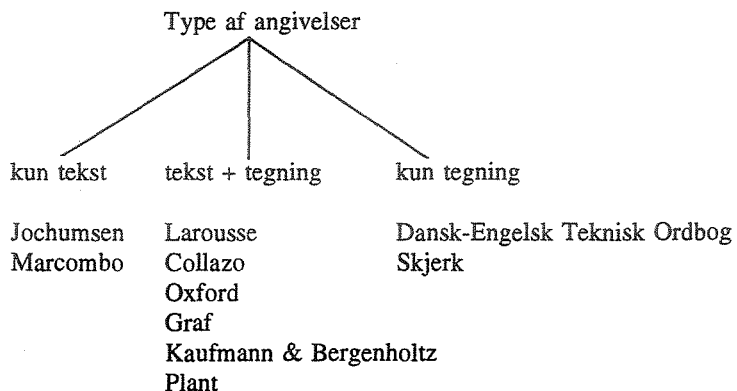
Skema 4:



Warrern og Clausen er de mest almindelige tekniske ordbøger til og fra dansk (med henholdsvis engelsk, tysk og fransk som fremmedsprog). Her er der altså overhovedet ikke medtaget encyklopædiske oplysninger. De tre ordbøger, der medtager enkelte oplysninger af encyklopædisk art, er også, hvad man traditionelt forstår ved oversættelsesordbøger. Ser man derimod på de ordbøger, som generelt inkluderer disse angivelser, drejer det sig næsten alle om en- og tosprogede ordbøger, som kan medregnes til klassen af deciderede encyklopædiske ordbøger. Bortset fra Kaufmann & Bergenholtz, der er et resultat af en nyudvikling, og som medtager oplysninger af både encyklopædisk og sproglig (grammatisk) art, så stiller kun Collazo sig bevidst denne opgave, idet den dog i praksis er temmelig karrig med hensyn til rent sproglige angivelser. Konklusionen må således blive, at en succesrig kombination af sproglige og encyklopædiske oplysninger langt fra er et generelt fænomen i tekniske ordbøger. Og fremfor alt må det konkluderes, at ordbøger, der decideret er beregnet på oversættelse, kun i meget lille omfang indtil videre har behandlet den encyklopædiske side af sagen.

Ser man på, hvilken type af angivelser de enkelte ordbøger anvender, får man et billede som i skema 5.

Skema 5:



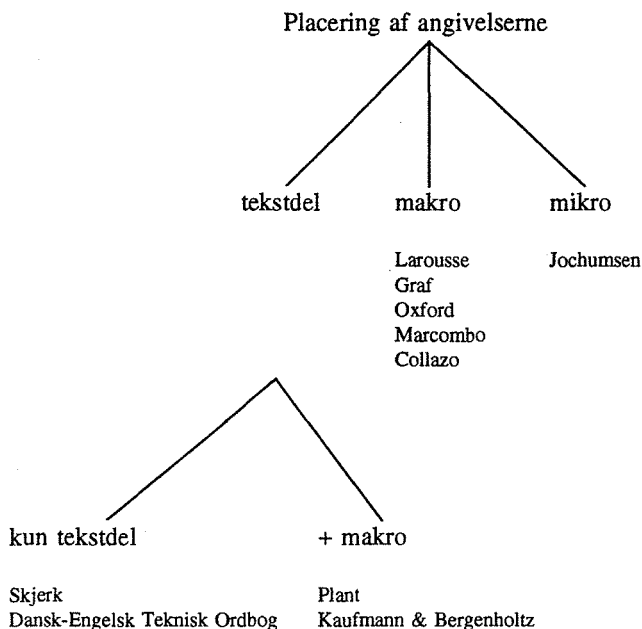
Heraf fremgår det, at over halvdelen af de ordbøger, som i større eller mindre omfang serverer encyklopædiske angivelser, har valgt en kombination af tekst og tegning. Af de ordbøger, som generelt medtager encyklopædiske oplysninger, har seks ud af syv denne kombination, idet Plant dog kun har enkelte tegninger, som er anbragt i indledningen. Med hensyn til de to ordbøger, som kun anvender tegning, er det henholdsvis en brancheordbog (Skjerk 1974) og en 40 år gammel teknisk ordbog, som endvidere har den helt særlige egenskab, at den undertiden angiver eksempler på anvendelse af de tekniske termer.

Hvad angår teksterne er der i det mindste to slags problemer. Det første er det sprog, hvorpå teksten er skrevet. Flere af disse ordbøger henvender sig til brugere på to sprog, hvorimod definitionerne eller de encyklopædiske angivelser kun er skrevet på det ene. Det vanskeliggør i høj grad forståelsen af et i forvejen svært tilgængeligt emne at skulle sætte sig ind i det på et fremmedsprog. Desuden er mange af forklaringerne skrevet i et sprog, som om det var videnskabsmænd eller teknikere, der henvendte sig til kolleger og ikke til lægfolk, således som ordbøgerne hævder, at de gør. Jochumsen (1983) er her værd at fremhæve, idet den såvel er skrevet på et letlæseligt sprog som på både dansk og spansk. Dog er de encyklopædiske oplysninger kun angivet sporadisk. Af disse observationer kan man drage to konklusioner, som ved første øjekast virker selvindlysende, men som åbenbart endnu ikke har fået et generelt praktisk udslag: 1. teksten skal så vidt muligt skrives på brugerens modersmål; 2. teksten skal forfattes i et for lægmand tilgængeligt sprog uden af den grund at blive fagligt udvandet, evt. kunne man bruge krydshenvisninger til at forklare vanskelige ord, som ikke kan undgås, og man bør også i det fornødne omfang anvende tegninger til at lette forståelsen.

Ser man på de ordbøger, der i større eller mindre grad bruger tegninger, vil man finde, at disse bruges på forskellig vis. Når der ses bort fra tegninger, som bruges af rent æstetiske årsager, kan man opstille tre hovedmåder at anvende tegningerne på: 1. tegningen illustrerer teksten, 2. teksten forklarer tegningen, 3. tekst og tegning overlapper hinanden. Selv om man kan finde mange eksempler på en uhensigtsmæssig kombination, har det ikke været muligt at opstille generelle regler for, hvornår den ene eller anden måde er den mest optimale, idet det må formodes, at dette spørgsmål må afgøres i det enkelte tilfælde. Man kan dog konstatere, at tilfælde af overlappning af tekst og tegning ofte har den meget nyttige funktion, at de på samme tid henvender sig til folk, der er enten visuelt eller tekstmæssigt disponerede. Derimod er der visse problemer i forbindelse med kun at bringe tegninger, uden forklaring. I det omfang, det ikke drejer sig om illustration af kendte genstande eller fænomener, siger tegningen ikke noget i sig selv, da den ikke på nogen som helst måde refererer til foreteelser i verden, som ordbogsbrugeren kan genkende. Her adskiller fagsprog sig klart fra almensprog, der som allerede nævnt henviser til ting, der indgår i de fleste menneskers almenviden. Lægmand er netop kendetegnet ved at være helt blottet for indsigt i det pågældende fag og dermed for evnen til at forbinde noget med og trække viden ud af tegninger uden den behørig forklaring. Undtaget herfra er illustration af farver, form og andre former for udseende, der ofte kun (eller i det mindste meget lettere) kan forklares med en tegning.

Det næste spørgsmål, jeg har undersøgt, er, hvordan de encyklopædiske angivelser er placeret i ordbogen. Resultatet af undersøgelsen ser ud som i skema 6.

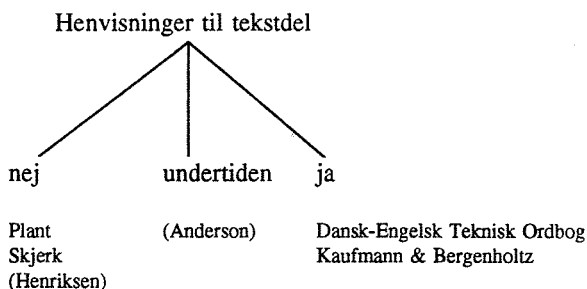
Skema 6:



Opdelingen i mikro- og makrostruktur samt tekstdel følger den typologi, som Hausmann og Wiegand (1990) opstiller. At oplysningerne er placeret i mikrostrukturen, betyder således, at man kun kan finde oplysninger inde i den enkelte ordbogsartikel. At de er placeret i makrostrukturen, forudsætter endvidere, at der er brugt krydshenvisninger, således at en del af den encyklopædiske viden om et enkelt ord gives i andre ordbogsartikler.

Af denne undersøgelse fremgår det, at det mest hyppige er at placere de encyklopædiske oplysninger i de enkelte ordbogsartikler og i større eller mindre omfang foretage krydshenvisninger mellem disse artikler. Med hensyn til placering i den overordnede tekstdel, dvs. i for- eller bagteksten, ser det endnu ikke ud til at være en almen praksis, i det mindste ikke når det drejer sig om tekst. Fælles for de to ordbøger, der kun har placeret oplysningerne her, er, at det i begge tilfælde drejer sig om tegninger, altså ingen tekst. Men de to ordbøger, der her er mest interessante ud fra et leksikografisk synspunkt, er Plant og Kaufmann & Bergenholtz. De har begge foretaget krydshenvisninger i makrostrukturen og har desuden placeret oplysninger i en såkaldt faglig indledning på henholdsvis 5 og ca. 50 sider, hvor der anvendes både tekst og tegninger. Ordbøgerne er begge decideret branche- eller fagorienteret (mikroelektronik og molekylærbiologi), og indledningen har hos begge den funktion at give en overordnet fremstilling af det pågældende fag. En sådan indledning kan selvsagt have mange fordele med hensyn til at give den nødvendige mængde encyklopædiske informationer. For at få maksimal gevinst heraf er der imidlertid brug for et system af henvisninger mellem selve ordbogsdelen og denne indledning. Jeg har derfor undersøgt, i hvor høj grad de ordbøger, som bringer oplysninger i tekstdelen, har et sådant system af henvisninger. For at få nok oplysninger har jeg her medregnet to juridiske ordbøger, nemlig Anderson og Henriksen (se skema 7).

Skema 7:



Kun to af ordbøgerne har altså generelt henvisninger ud til tekstdelen, idet jeg hos Anderson kun har fundet ialt to eksempler herpå. I både Plant, Anderson, Henriksen og Kaufmann & Bergenholtz drejer de encyklopædiske informationer i tekstdelen sig om en faglig indledning. Man må imidlertid konstatere, at kun i Kaufmann & Bergenholtz får man det fulde udbytte af denne indledning, idet der er foretaget en omfattende henvisning fra artiklerne ud til indledningen og tilmed også den anden vej. Dog har Henriksen (som i øvrigt også Kaufmann & Bergenholtz) udarbejdet en fyldig stikordsliste til indledningen, så man forholdsvis let kan gå til den. Omvendt svæver indledningen hos Plant og Anderson ligesom i luften, uden at man rigtig kan bruge den til noget, med mindre man hver gang vil læse den i sin helhed eller skimme sig frem til den ønskede oplysning (hvis den da er der).

Med hensyn til henvisninger er der endnu et problem, som gælder for både henvisninger internt i makrostrukturen og henvisninger ud i tekstdelen. Det er spørgsmålet om antallet af henvisninger og deres slutadresse. Alt for ofte er der et meget stort antal henvisninger, som besværliggør læsningen og navnlig forståelsen af teksten. F.eks. har jeg i den spanske udgave af Theodor Lewanskis »Diccionario de Lingüística« talt flere dusin henvisninger i en enkelt artikel. Under »sintaxis« er der således 32. Desuden anvender samme ordbog ligesom mange andre, hvad man kunne kalde cirkelhenvisninger: man bliver i sin søgen efter at forstå et bestemt begreb henvist fra den ene artikel til den anden for til sidst at ende, hvor man begyndte, og ofte med samme indsigt, som man startede med. Disse konstateringer fordrer, at der foretages mere systematiske undersøgelser af henvisningernes rolle i fagordbøger, og at der opstilles vejledende principper for deres anvendelse. Dog kan man allerede her konstatere, at et af problemerne ikke i sig selv består i at reducere antallet af henvisninger, men i at skrive teksten i et for lægmand mere forståeligt sprog, hvormed henvisningerne i et vist omfang overflødiggøres.

Men henvisninger bør ikke af den grund betragtes som et onde. Tværtimod spiller de en yderst nyttig rolle, idet de rigtigt brugt så at sige binder alle de små adskilte tekststykker i en ordbog sammen til et hele, hvormed brugeren bliver i stand til at vride langt flere encyklopædiske informationer ud af ordbogen. Udviklingen i retning af ikke blot at foretage krydshenvisninger i makrostrukturen, men også at udarbejde deciderede faglige indledninger, der gennem henvisninger og stikordsregistre indgår i ordbogen som et hele og bringer den op på et højere plan (i et vist omfang hos Henriksen og mere udviklet hos Kaufmann & Bergenholtz), er en udvikling, der bestemt må hilses velkommen.

Som konklusion på denne lille undersøgelse kan man sige, at en række tekniske ordbøger mere eller mindre bevidst synes at have indset nødvendigheden af at medtage forskellige former for encyklopædiske informationer, men at måden at præsentere dem på i

disse ordbøger er stærkt varierende og ofte virker noget tilfældig. Desuden udnytter de sjældent alle de redskaber, der findes til at give brugeren maksimal glæde af disse oplysninger.

Vi er her ved et punkt, hvor den teoretiske leksikografi ikke længere kan nøjes med at være beskuende og undersøge de eksisterende ordbøger. I sig selv er det ikke dårligt at undersøge resultatet af andre menneskers bevidste arbejdsindsats. Man kan hente megen inspiration derfra. Men en leksikografi, der lægger for megen vægt på dette, risikerer let at blive for traditionsbundet og overse vigtige problemstillinger, hvormed den nødvendige nytænkning udebliver. Den teoretiske leksikografi må hæve sig op til rang af virkelig teori og anvise nye løsninger, idet den naturligt bygger på de mest fremskredne erfaringer. En stadig og frugtbar vekselvirkning mellem teori og praksis er vejen frem. Som en kendt og for tiden noget udskaældt herre engang sagde, gælder det ikke så meget om at fortolke verden som at forandre den.

## Litteratur

### 1. Ordbøger

- Anderson, R.B.J.: *Anglo Scandinavian Law Dictionary*, Universitetsforlaget, Oslo 1977  
*Clausens Tekniske Ordbøger Dansk-Engelsk*, Grafisk Forlag 1990  
*Clausens Tekniske Ordbøger Dansk-Fransk*, Ashehoug Dansk Forlag 1989  
*Clausens Tekniske Ordbøger Dansk-Tysk*, Grafisk Forlag 1990  
*Clausens Tekniske Ordbøger Fransk-Dansk*, Grafisk Forlag 1990  
Collazo, Javier L.: *Diccionario Enciclopédico de Términos Técnicos, inglés-español, español-inglés*, McGraw-Hill Book Company, Madrid 1980  
*Dansk-Engelsk Teknisk Ordbog*, Einar Hacks Forlag, København 1950  
*Diccionario de Términos Científicos y Técnicos*, Marcombo Boixareu Editores, Barcelona 1981  
*Diccionario Ilustrado de las Ciencias*, Larousse, Paris 1987  
*Diccionario Oxford de Informática, inglés-español, español-inglés*, Ediciones Diaz de Santos, Madrid 1985  
*Enciclopedia de la Técnica y de la Mecánica*, Ediciones Nauta, Barcelona 1973  
Galindo, Orlando: *Dansk Spansk Fagordbog*, Teknisk Forlag, København 1979  
Graf, Rudolf F.: *Diccionario de Electrónica*, Ediciones Pirámide, Madrid 1988  
Henriksen, Torben: *Spansk Dansk Juridisk Ordbog*, Gad, København 1991  
Jochumsen, Svend: *Teknisk Ordbog Dansk Spansk*, Gyldendal, København 1983  
Kaufmann, Uwe & Bergenholtz, Henning: *Genteknologisk leksikon. Dansk-engelsk og engelsk-dansk ordbog over molekylærbiologi og DNA-teknologi*, udkommer 1992  
Lewandowski, Theodor: *Diccionario de Lingüística*, Catedra. Madrid 1986  
Mink, H.: *Diccionario Técnico Alemán-Español*, Herder, Barcelona 1987  
Mink, H.: *Diccionario Técnico Español-Alemán*, Herder, Barcelona 1985  
Plant, M.: *Diccionario de Microelectrónica*, Paraninfo, Madrid 1987  
Skjerck, E.: *Engelsk / Dansk Bilteknisk Ordbog*, Lentz & Jenssens Forlag, Ringsted 1974  
Warrern: *Dansk-Engelsk Teknisk Ordbog*, Clausens Bøger 1988  
Warrern: *Dansk-Tysk Teknisk Ordbog*, Clausens Bøger 1984  
Warrern: *Engelsk-Dansk Teknisk Ordbog*, Clausens bøger 1981

## 2. Anden litteratur

- Duvå, Grete / Laursen, Anna-Lise / Maidahl, Lisbet (1992): Brugerundersøgelse vedrørende oversættelse af fagtekst. I: Ruth Vatvedt Fjeld (red.): *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.-31.mai 1991*. Oslo.
- Eco, Umberto (1984): Metaphor, Dictionary, and Encyclopedia. I *New Literary History* 15, 255-271.
- Haiman, John (1980): Dictionaries and Encyclopedias. I: *Lingua* 50, 329-357.
- Hausmann, Franz Josef (1985): Lexikographie. I: Christoph Schwarze / Dieter Wunderlich (Hrsg.): *Handbuch der Lexicologie*. Königstein: Athenaum, 367-411.
- Hausmann, Franz Josef / Wiegand, Herbert Ernst (1990) Component Parts and Structures of General Monolingual Dictionaries: A Survey. I: Franz Josef Hausmann / Oskar Reichmann / Herbert Ernst Wiegand / Ladislav Zgusta (eds.): *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography, First Volume*. Berlin, New York: Walter de Gruyter, 328-360. (=Handbücher zur Sprach- und Kommunikationswissenschaft, band 5.1.)
- Hermanns, Frits (1988): Das Lexikographische Beispiel. Ein Beitrag zu seiner Theorie. I: Gisela Harras (ed.): *Das Wörterbuch. Artikel und Verweisstrukturen*. Düsseldorf: Schwann / Bagel, 161-195. (=Sprache der Gegenwart 74)
- Hupka, Werner (1989): Das enzyklopädische Wörterbuch. I: Franz Josef Hausmann / Oskar Reichmann / Herbert Ernst Wiegand / Ladislav Zgusta (eds.): *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography, First Volume*. Berlin, New York: Walter de Gruyter, 988-999. (=Handbücher zur Sprach- und Kommunikationswissenschaft, band 5.1.)
- Hupka, Werner (1989a): Wort und Bild. Die Illustrationen in Wörterbüchern und Enzyklopädiën. Tübingen: Max Niemeyer Verlag. (=Lexicographica, series maior 22)
- Kalverkämper, Hartwig (1990): Das Fachwörterbuch für den Laien. I: Franz Josef Hausmann / Oskar Reichmann / Herbert Ernst Wiegand / Ladislav Zgusta (eds.): *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography, Second Volume*. Berlin, New York: Walter de Gruyter, 1512-1523. (=Handbücher zur Sprach- und Kommunikationswissenschaft, band 5.2.)
- Mikkelsen, Hans Kristian (1991): Towards the ideal special language translation dictionary. I: *Hermes* 6, 91-109.
- Scholfield, P.J. (1980): Evaluation Selection Policy and Grammatical and Semantic Information in an EST Dictionary. I: *Fachsprache*, bd. 2, 98-108.
- Wiegand, Herbert Ernst (1988): Was eigentlich ist Fachlexikographie?. I: Horst Haider Munske / Peter von Polenz / Oskar Reichmann / Reiner Hildebrandt (eds.): *Deutscher Wortschatz. Lexikologische Studien. Ludwig Erich Schmitt zum 80. Geburtstag von seinen Marburger Schülern*. Berlin, New York: Walter de Gruyter, 729-790.

## En type bevegelsesverb i norsk, og hvordan de er behandlet i noen norsk-franske ordbøker

Norsk inneholder en rekke verb som uttrykker både bevegelse og måte. Disse verbene oppfører seg i visse kontekster helt forskjellig fra bevegelsesverb i fransk. Dette bør en norsk-fransk ordbok ta hensyn til. I det følgende blir det undersøkt i hvilken grad dette har blitt gjort i eksisterende ordbøker, og det blir presentert visse momenter som kan spille en rolle for en adekvat og systematisk presentasjon av slike verb i fremtidige ordbøker.

Beskrivende bevegelsesverb, dvs. verb som samtidig kan uttrykke en forflytning fra et sted til et annet og måten det skjer på, er svært hyppige i norsk. Her er noen tilfeldige eksempler fra reportasjer i Aftenposten i år:

Publikum karrer seg oppover bakkene til benkene og anlegget.  
(19. juli, 40)

Når hun har svinset vekk bak teppet, ja, da dukker Per Asplin frem.  
(19. juli, 40)

Og så smådanser hun av scenen i en kjole som ble ferdigsydd igår.  
(19. juli, 40)

- Du er jo bare såvidt rød, fnyser hun idet vi matt sjangler mot verandadøren.  
(19. juli, 40)

Så myldrer vi inn i parken der japanerne straks begynner å fotografere.  
(19. juli, 38)

Når det norske sommerværet kommer sigende inn fra Nordmarka med en isende vind og en liten regnskur, trekker resten inn i souvenirbutikken.  
(19. juli, 38)

Sykkelturene over den sørvestlige delen av Hardangervidda er blitt en kjempesuksess. NSB regner med at over 6000 personer vil trække seg over vidda i løpet av sommeren.  
(21. juli, 1)



I tillegg til fly har jeg en trikk (hyggelig) som skrangler forbi hvert 20. minutt, biler selvfølgelig, og av og til lyder fra NSB's lokaltog som suser forbi nede på Skøyen.

(25. juli,12)

Siden vårt språk er så rikt på slike verb, er det selvfølgelig uhyre viktig å gi dem en fylldig behandling i aktive tospråklige ordbøker. Når det gjelder forholdet mellom norsk og fransk, er det slett ikke sikkert at fransk har like mange beskrivende bevegelsesverb som norsk har, og resultatet blir at den samme oversettelsen til fransk må brukes flere ganger. Bare tenk på ord som *tasse, tasle, tusle, tutle, pusle; skumpe, skumple; krype, kravle*. Når det gjelder en tospråklig norsk-fransk ordbok er det imidlertid ikke nok å finne fram til en god oversettelsesekvivalent for selve verbet. Disse beskrivende bevegelsesverbene står som oftest sammen med et stedsadverbial, en retningsangivelse. I hvert av eksemplene ovenfor finner vi et slikt adverbial. Det siste eksemplet inneholder to verb, *skrangle* og *suse*, som alene ikke uttrykker noen bevegelse. Sammen med et retningsadverbial blir de imidlertid beskrivende bevegelsesverb. Fenomenet virker særlig å gjelde verb som beskriver en lyd, og ser ut til å kunne brukes i alle mulige kombinasjoner. Man har selvfølgelig ikke så store sjanser til å støte på andre enn de mest alminnelige av dem i en ordbok. Et annet karakteristisk trekk ved norsk er evnen til å lage nye ord ved hjelp av sammensetninger, (jf. *smådanse* ovenfor). Her vil det selvsagt også bli lakuner. Begge disse fenomenene blir godt illustrert av dette lille avsnittet fra Jon Michelet: *Den gule djevelens by*:

Bilen svingte på to hjul inn i den lange Columbia Street, ulte inn under brusystemet der Brooklyn-Queens Expressway møtte Gowanus Expressway, kvein seg fram mellom biler som hadde stoppet i krysset og råflådde nedover mot havna. Kojak og Cannon ville ha kost seg, mens Colombo ville vært noe betenkt.

(125)

(Hverken *ule inn under, kvine seg fram* eller *råflå* finnes i ordbøker.)

Fransk har selvfølgelig også en del beskrivende bevegelsesverb, men de kan bare kombineres med få typer av retningsangivelser. Litt forenklet kan man si at fransk har et sett med verb som er beskrivende, og et sett med verb som uttrykker retning. Man kan fremstille det skjematisk på denne måten:

	bevegelse	måte	retning
norsk	+	+	+
fransk a)	+	+	-
fransk b)	+	-	+

Under fransk a) finnes *courir* (løpe), *se traîner* (slepe seg), *ramper* (krype), *grimper* (klatre)... Under b) *monter, descendre, entrer, sortir, passer, traverser, avancer...*

Retningene som uttrykkes i b) er opp(over), ned(over), inn (i), ut (av), forbi, over, fremover... Måten uttrykkes ved et tillegg, som kan bestå av det opprinnelige bevegelses verbet

i gerundium, som f.eks. i *monter l'escalier en courant* (løpe opp trappen), et adverb som i *s'introduire furtivement dans sa chambre* (lure seg inn på rommet hans), eller et preposisjons-syntagma som i *rentrer à la nage* (svømme inn). Dette kontrastive aspektet må også komme fram i ordboksartikkelen. På dette området er den tospråklige ordboken den enspråklige totalt overlegen. Slår man opp på *courir* f.eks. i en fransk-fransk ordbok, vil man ikke finne noe som kan hjelpe en til å oversette *løpe opp/ned/inn/ut* osv. Man vil være like hjelpeløs hvis man slår opp på et av retningsverbene og håper å komme seg videre derfra.

Hvis man skal undersøke hvordan disse norske verbene er behandlet i norsk-franske ordbøker av nyere dato, er det bare to å forholde seg til. Det er *Kunnskapsforlagets Norsk-fransk ordbok*, utkommet i revidert utgave i 1990, og Universitetsforlagets *Stor norsk-fransk ordbok* som kommer ut for første gang i 1991. Siden det er en ganske stor forskjell i størrelse på de to ordbøkene, er det egentlig uhensiktsmessig å sammenligne dem, så jeg vil nøye meg med å beskrive og kommentere hvordan disse verbene er behandlet i *Stor norsk-fransk ordbok*, som jeg også selv har vært med på å utarbeide.

Jeg vil først se om denne ordboken er i stand til å løse oversettelsesnøttene som finnes i eksemplene tatt fra *Aftenposten*. Det viser seg at alle verbene finnes som oppslagsord bortsett fra *karre seg* og *smådanse*. *Smådanse* er ikke med fordi det er en av utallige sammensetninger. *Karre seg* er derimot et forholdsvis frekvent verb i norsk og burde absolutt vært med. *Sjangle* har fått tynnest dekning:

**sjangle** tituber, chanceler

*Svinse* og *trække* er fyldigere, men har ikke fått eksempel på kontekst med typen stedsadverbial som krever omskrivning:

**svinse** (*vrikke*) tortiller, agiter, remuer: **s- med halen** remuer (*el agiter*) la queue; **s- med hoftene** tortiller des hanches; **s- hit og dit** (*el omkring*), **s- og svanse** s'affaires (*el valser*) de-ci de-là (*el en tous sens, de droite et de gauche*)

**trække** piétiner: **gå og t- dagen lang** aller et venir toute la journée; **t- noe i stykker** piétiner qch; **t- noe ned i sneen** enfouir qch dans la neige; (*fig*) **t- noen ned i sølen** traîner qn dans la boue; **t-opp en løype** tracer une piste; **trakk ikke på plenen** (*oppslag*) pelouse interdite; **t- noen på tærne** marcher sur les pieds de qn

*Myldre*, *sige* og *skrangle* har alle med minst ett eksempel som viser tydelig kontrasten mellom norsk og fransk:

**myldre** fourmiller, grouiller: **det myldret av folk i gatene** les rues fourmillaient (*el grouillaient*) de monde; **folk myldret ut av fotballstadion** les gens sortaient en foule du stade

**sige** (1) suinter: **det siger vann gjennom muren** l'eau suinte à travers le mur; **karene seig inn etter hvert** les hommes sont entrés peu à peu; **motløsheten seig inn på ham** le découragement commençait à le gagner (*el s'insinuait en lui*); **skipet seig inn til brygga** le navire est lentement venu à quai (*el a lentement accosté*)

**skrangle** faire un bruit de ferraille: **skjermen skrangler** le gardeboue fait un bruit de ferraille; **bilen skranglet forbi** la voiture est passée avec un bruit de ferraille; **det skrangler i boksen** qch fait du bruit dans la boîte; **s- med pengene i lommen** faire tinter la monnaie dans sa poche; **vi satt og skranglet på toget** on était secoués (*el ballottés*) dans le train; **vi skranglet sammen til en flaske vin** nous nous sommes (*el nous avons*) cotisé(s) pour une bouteille de vin

*Suse* er det verbet som har fått den mest systematiske behandlingen. Der har man satt opp de mest vanlige preposisjonene *suse* følges av, og dermed fått med de forskjellige oversettelsesvariantene. Dessuten har konstruksjonen *komme + presens partisipp* av bevegelses verbet kommet med. Denne typen er karakteristisk for mange bevegelsesverb, og bør derfor finnes:

**suse** (2) (*bevege seg hurtig*) **det gikk så det suste** ça filait (*el fonçait, gazait, allait à fond de train*); **s- av gårde/forbi/nedover** partir/passér/descendre à toute vitesse (*el à toute allure, en trombe, à fond de train*); **la humla suse** s'en fiche, (*fam*) s'en foutre, laisser pisser le mérinos; **i susende fart** à toute vitesse (*el allure*), (*med kjøretøy*) à pleins gaz; **komme susende** arriver à toute vitesse (*el à toute allure, en trombe, à fond de train*)

Etter en del stikkprøver tror jeg jeg kan si at disse eksemplene er representative for behandlingen av beskrivende bevegelsesverb i Stor norsk-fransk ordbok. De færreste består som *sjangle* av oversettelsekvivalenten alene. Men man kan finne noen eksempler på det, f.eks. *kreke, jogge, lunte*:

**kreke** se traîner péniblement

**jogge** faire du jogging

**lunte** (*om mennesker*) marcher lentement, (*om dyr*) aller au petit trot

Som man vil ha skjønt, har leseren ut fra disse ordboksartiklene ikke noen særlig mulighet til å lage setninger som: *Gamlingen kreket seg opp bakken, De jogget til butikken, De luntet nedover bakken*, o.l. De fleste har imidlertid med et eksempel på omforming ved tilføyelse av retningsadverbial. Jeg vil ta med et par oppslagsord til for å vise dette:

**ramle** (1): (*larve*) **r- nedover trappene** descendre l'escalier à grand bruit; **vognen ramlet over broen** la charette traversa bruyamment le pont

**toget** (2) (*gå i flokk*): **de toget forbi** ils sont passés en cortège; **de kom togende inn på plassen** ils sont arrivés en cortège sur la place

Jeg synes resultatet av denne undersøkelsen er relativt tilfredsstillende, men man kan jo ikke si at behandlingen av disse verbene har noen virkelig koherens. Ett problem med denne typen verb virker også å ha blitt oversett. Ved noen retningsadverbialer kan man bruke det beskrivende bevegelses verbet uten noen omskriving med retningsverb. Det gjelder særlig adverbialer innledet med *vers* og *jusque* (*mot, til*). Ingenting forteller brukeren av ordboka at

det er slik. Siden det såvidt jeg vet ikke er noe eksempel på dette i ordboka, vil jeg illustrere det med et par eksempler fra oversatte romaner, en norsk oversatt til fransk (Knut Faldbakken: *Bad Boy*, oversatt av Eric Eydoux), og en fransk oversatt til norsk (Benoîte Groult: *Les Vaisseaux du coeur*, oversatt av Kari Kemény),

Mens jeg løper mot bussholdeplassen,  
(F, bb, 70)

Tout en courant vers l'arrêt d'autobus,  
(70)

Il a posé sa tempe contre la mienne d'un geste familier et nous avons roulé jusqu'au parking souterrain sans un mot.  
(G, v, 179)

Han la tinningen mot min på sin velkjente måte, og vi trillet ned til den underjordiske parkeringsplassen uten et ord.  
(116)

la jeune fille folle qui courait vers l'île, vers les îles du bout du monde,  
(G, v, 211)

den unge, gale jenta som løp mot øya, mot øyene på verdens ende,  
(138)

En koherent beskrivelse bør altså inneholde fire punkter: oversettelse uten retningsadverbial, oversettelse med et knippe retningsadverbialer, oversettelse med *vers*, *jusque...*, oversettelse av *komme + presens partisipp*.

De oversettelsene som er vist fra Stor norsk-fransk ordbok må kunne kalles maksimale oversettelser, dvs. man har villet oversette hele betydningsinnholdet. Det står da igjen å se om franskmenn alltid bruker de parafasene ordboka foreslår. Én måte å gjøre det på er å undersøke oversettelser av norske romaner som ofte kryr av beskrivende bevegelsesverb. Her vil jeg igjen bruke oversettelsen av *Bad Boy* som eksempel:

Strandvaktene løp opp og nedover stripen som hørte til hotellet,  
(F, bb, 54)

Les surveillants remontaient et descendaient en courant la plage de l'hôtel,  
(55)

Og jeg snublet avgårde midt i videofilmen for å rable ned denne beskjeden til meg selv.  
(F, bb, 68)

Sur quoi, abandonnant mon film vidéo, je suis parti d'un pas chancelant pour aller gigouiller ce message à mon intention.  
(68)

Her har oversetteren brukt de mulighetene jeg har nevnt tidligere: *retningsverb* + *gerundium* av *bevegelsesverbet* eller *måtesadverbial*. I andre tilfeller nøyer han seg med *retningsverbet*:

Den samme, ødeleggende latteren jeg følte boble i brystet da jeg løp ned trappene klokken fem på ett igår,

(F, bb, 25)

Ce même rire destructeur que je sentais bouillonner dans ma poitrine lorsque, les paumes brûlantes et le poing égratigné, je descendais l'escalier à une heure moins cinq,

(27)

og løp, ja løp gjennom korridoren, ned trappen, ut av den tunge, slitte hoveddøren, over skolegården, ut av porten.

(F, bb, 61)

et me mis à courir, oui à *courir* à travers le corridor, pour ensuite descendre l'escalier, sortir par la lourde porte principale puis traverser la cour de l'école et enfin passer le porche.

(62)

Det går vel klart fram av det siste eksemplet hvorfor oversetteren ikke har føyet til *en courant* først til *descendre*, så til *sortir* og *traverser* for å avslutte med *passer le porche en courant*. Det ville stilistisk sett bli en helt umulig lang og klønete setning. Dessuten har han et *courir* i starten, som jo antyder at personen fortsetter å løpe. Men hvor elegant og økonomisk kan ikke dette uttrykkes på norsk!

I det første eksemplet er motivasjonen for å prioritere det stilistiske fremfor meningen mindre tydelig. Men oversetteren har nok synes at de innskuddene som allerede finnes i setningen gjør den tilstrekkelig lang, uten at man også skal føye til et *en courant*. Men noe av meningen går altså tapt.

Det franske språk har imidlertid andre virkemidler som kan vise på hvilken måte en forflytning skjer. I det følgende kjente eksempel av Vinay og Darbelnet er det subjektet som tydeliggjør at man *rir* inn i gården, og det er tilstrekkelig å bruke et *retningsverb*. (*Cavalier* betyr *rytter*):

Les cavaliers sont entrés dans la cour.

Andre ledd i en setning kan anskueliggjøre det samme. Jeg vil igjen bruke et eksempel fra Groult:

Ils ont loué des bicyclettes et font le tour de l'île chaotique jusqu'aux blocs erratiques de granit...

(G,v,108)

Siden *sykler* blir nevnt først i setningen, er det helt tilstrekkelig å bruke et *retningsverb* alene. Det kan være morsomt å se hvordan denne setningen har blitt oversatt til norsk:

De har leid sykler, og trækker rundt den kaotiske øya helt til vandreblokkene av polert granitt...

(77)

Oversetteren har helt korrekt brukt et beskrivende bevegelsesverb. At ordet *sykler* allerede står der, influerer ikke på valget av verb. Man kunne vel brukt verbet *sykle* også uten at det hadde virket rart.

Det faktum at den maksimale oversettelsen ikke alltid blir brukt, skaper selvfølgelig et problem for ordboksforfatteren. Siden alle leksikalske og grammatikalske opplysninger i ordboka finnes implisitt i eksemplene, er det vanskelig å gjøre leseren oppmerksom på problemet. Kanskje det til kravet om koherens også bør føyes til et femte punkt, nemlig at det regelmessig skal være med et eksempel av denne typen? De finnes allerede i Stor norsk fransk ordbok, men selvfølgelig bare de tydeligste, som under oppslagsordet *dampe*:

**båten dampet oppover elven le bateau remontait la rivière**

Som avslutning har jeg lyst til å se på hvordan noen franske retningsverb er behandlet i en fransk-norsk ordbok, nemlig Kunnskapsforlagets som kom ut i helt omarbeidet utgave i 1986. Under *traverser* finner man bl.a.: *gå (komme, løpe, renne, kjøre, seile, reise, fly etc.) over el. gjennom*. Under *descendre* *gå (komme, stige, løpe, kjøre osv.) ned*. Leseren blir altså gjort oppmerksom på at disse verbene må oversettes med et eller annet beskrivende bevegelsesverb. Sammenlikner man med en tidligere utgave, den fra 1976, vil man der finne for *traverser* bare *gå tvers igjennom, over, og for descendre gå ovenfra og nedover*.

### Sitert litteratur

Aftenposten

Faldbakken, Knut. 1990. *Bad Boy*. (Utg. Gyldendal Pocket.) Oslo.

Faldbakken, Knut. 1990. *Le Monarque*. (Overs. Eric Eydoux.) Paris.

Groult, Benoîte. 1988. *Les vaisseaux du coeur*. Paris.

Groult, Benoîte. 1990. *Hjertets kar*. (Overs. Kari Kemény). Oslo.

Michelet, Jon. 1991. *Den gule djevelens by*. (Utg. Den norske Bokklubben.) Oslo.

### Ordbøker

Dedichen, Line. 1976. *Fransk-Norsk ordbok*. Oslo.

Elligers, Anne. 1986. *Fransk-norsk ordbok*. Oslo.

Grundt, Lars Otto (hovedred.). Grundt, Micheline. Jacobsen, Tove. 1991. *Stor norsk-fransk ordbok*. Oslo.

**Faglitteratur**

- Jacobsen, Tove. 1987. Eksemplene i to-språklige ordbøker. I: Tove Jacobsen/Jan Roald red. *Flerspråklig leksikografi og terminologi. Seminarrapport. Bergen 22.-23. mai 1987*: 77-83. Bergen.
- Jacobsen, Tove. 1989. Oversettelse eller utelatelse av visse måtesadverb i tospråkligeordbøker. I: *TRIBUNE. Skriftserie for Romansk institutt, Universitetet i Bergen*, 8: 44-52. Bergen.
- Snell-Hornby, Mary. 1983. *Verb-descriptivity in German and English*. Heidelberg.
- Snell-Hornby, Mary. 1987. Towards a Learner's bilingual dictionary. I: Anthony Cowie ed. *The Dictionary and the Language Learner. Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985*. Tübingen
- Vinay J.P., Darbelnet J. 1958. *Stylistique comparée du français et de l'anglais*. Paris.

## Norsk EF-terminologi

Om det pågående arbeidet med å registrere norsk terminologi knyttet til oversettelsen av deler av EFs regelverk til norsk

I forbindelse med de pågående EØS-forhandlingene (sv. EES) mellom EFTA og EF (sv. EG) er det i de nordiske EFTA-stater under oversettelse en stor mengde EF-rettsakter til nasjonalspråket. I Norge blir oppgaven ivarettatt av Oversettelsesenheten/EØS i Det kongelige utenriksdepartement. Terminologisk konsekvens er en nødvendighet i lovttekster for at de skal fungere i praksis. Oversettelsesarbeidet krever derfor en kraftig innsats også på det terminologiske plan. En del av termene i rettsaktene er problematiske å registrere. Ikke minst gjelder dette når det er semantisk-kognitive ulikheter mellom kilde- og målspråk, eller når termen er en konstruksjon for å tilfredsstille EFs lovgivningsbehov.

### Bakgrunn

Siden juni 1990 har det vært ført forhandlinger mellom EFTA og EF med sikte på å komme fram til en avtale om et «europæisk økonomisk samarbeidsområde», forkortet EØS. Forhandlingene skulle etter planen ha vært sluttført i juni 1991. I oktober har det imidlertid ennå ikke lyktes partene å bli enige på alle punkter.

I Norge skal en eventuell EØS-avtale snarest mulig legges fram for de folkevalgte som stortingsproposisjon. Dersom Stortinget godkjenner avtalen, vil også Norge være en del av EFs «indre marked» når dette etter planen realiseres f.o.m. 1. januar 1993.

### Oversettelse av relevante deler av EFs regelverk

De deler av EFs regelverk som skal regulere det indre marked, må innenfor rammen av en EØS-avtale også inngå i EFTA-statenes lovgivninger. EØS-forhandlingene har således i Norge og i de øvrige nordiske EFTA-stater skapt et behov for å få disse delene av regelverket oversatt til nasjonalspråket.

Antall dokumenter som skal oversettes er 1 200-1 400, med et gjennomsnittlig sidetall på ca. ti trykte sider med dobbel spaldebredde. Dette motsvarer minst 30 000 maskinskrevne A 4-sider.

Arbeidet med å oversette de avtalerelevante delene til norsk har to overordnede siktemål:

- å gi politiske og faglige miljøer, samt allmennheten, mulighet til å vurdere og ta stilling til avtalen før den endelige avgjørelse faller i Stortinget,
- å sikre at disse delene av regelverket foreligger i norsk språkdrakt når en eventuell EØS-avtale trer i kraft.



I Norge ble det høsten 1990 opprettet en midlertidig «oversettelsesenhets» i Utenriksdepartementet med oppgave å utføre dette arbeidet. Oversettelsesenheten/EØS omfatter i oktober 1991 totalt 17 heltidsansatte, fordelt på 1 byråsjef, 8 oversettere, 6 tekstrevisorer, 1 terminolog og 1 kontoransatt. I tillegg settes tekster ut til ca. 20 frilansoversettere.

### Originaltekstene

Dokumentene som oversettes, er *rettsakter*, dvs. juridiske tekster avledet av EFs "grunnlov", Romatraktaten. Den alt overveiende del er *direktiver*, men det forekommer også *forordninger*, *rådsvedtak* og *domstolsavgjørelser*. Alle er av bindende karakter. I tillegg kommer selve avtaleteksten med tilhørende vedlegg og protokoller (ca. 1 000 sider), og om og når en EØS-avtale trer i kraft, vil også en del ikke-bindende tekster måtte oversettes.

Forordninger og direktiver er imidlertid bindende på ulik måte og har ulik rettskraft. Forordningene er *direkte bindende for borgere og juridiske personer* i alle medlemsstater (og innen et ev. EØS). De er gjeldende lov innen hele Fellesskapet med nøyaktig den ordlyd de er vedtatt i.

Direktivene er *direkte bindende for medlemsstatene, men indirekte for borgere og juridiske personer*. I praksis innebærer dette at det ikke er rettsaktens ordlyd som er bindende, men dens intensjon. Til forskjell fra forordningene skal direktivene innarbeides, eller inkorporeres, i de nasjonale lovgivninger som lover og forskrifter. Når det gjelder direktivene, kan derfor tekstene i noen grad tilpasses hevdvunnen formuleringsskikk og språkbruk i de enkelte medlemsstater.

Oversettelsesarbeidet er krevende. Ettersom det er lovttekster som oversettes, skal det ideelt sett råde absolutt ekvivalens mellom original og oversettelse. Ulikheten mellom forordninger og direktiver får naturligvis følger for arbeidet; forordningene må være lytefritt oversatt med hensyn til både form og innhold, mens enkelte små avvik for direktivenes del ikke vil få de samme konsekvenser. Dette forholdet gjør seg også gjeldende i det terminologiske arbeidet. Termer i forordninger får sin endelige form ved oversettelsen. De må derfor være fullstendig ekvivalente både i intensjon (betydning) og ekstensjon (begrepsomfang). Når det gjelder termer fra direktiver, er kravet til ekvivalens ikke så absolutt, ettersom det vil være mulig å nasjonalisere eller endre/omdefinere en del termer når direktivene gjennomføres i nasjonal lovgivning. I praksis er likevel det terminologiske arbeidet like krevende for begge typer rettsakter, ettersom utgangspunktet hele tiden er å finne fram til rett term fra første stund.

### Terminologi

Samtlige oversettere ekserperer og registrerer termer i tekstene de oversetter. Terminologen har ansvar for organisering og samordning av terminologiarbeidet, samt for oppbygging av en elektronisk søkbar termbase, til daglig kalt «UD-basen». Basen vil inneholde mye av den sentrale EF-terminologi, men også alminnelig teknisk og annen terminologi brukt i regelverket.

Databasen byr på store fordeler framfor tradisjonelt ordboksformat. Den kan bl.a. tillates å bli stor uten å virke uoversiktlig, det kan søkes etter ulike typer informasjon og

søkene kan utføres på forskjellige måter. Programvaren som benyttes, er utarbeidet ved Norsk termbank i Bergen.

Termbasen er trespråklig, med både fransk og engelsk term som oppslagsord ved siden av den norske ekvivalenten. Den franske termen står først og er styrende, liksom den franskspråklige tekstversjonen går foran den engelske ved oversettelsen til norsk. Flerspråkligheten vil gagne framtidige brukere av basen, men er også til fordel under oversettelsesarbeidet, ettersom den gjør det lettere å identifisere rett norsk term. Iblant kan imidlertid flerspråkligheten være en ulempe, nemlig når det er mangelfullt samsvar mellom fransk og engelsk term, noe som ideelt sett ikke burde forekomme.

Det fins flere motiver for å bygge opp en termbase i tilknytning til oversettelsesarbeidet:

- Terminologisk støtte er til stor hjelp for oversetterne i deres daglige arbeid og bidrar dessuten til å gjøre oversettelsene mer enhetlige og konsekvente. Det sistnevnte er ikke minst viktig når man tar i betraktning sjangeren og den store tekstmengden.
- Oversettelse eller forståelse av tilsvarende EF/EØS-dokumenter i framtida vil falle lettere når viktige deler av terminologien allerede er kartlagt og dokumentert.
- En termbase av denne type vil ha stor verdi for ulike fagmiljøer i både offentlig og privat sektor, f.eks. ved markedsføring av produkter og varer eller ved utarbeidelse av norske versjoner av europeiske standarder o.l. Uansett utfallet av EØS-forhandlingene må man regne med at EFTA-statene vil måtte tilpasse seg EFs kommende harmoniserte indre marked, og tilpassningsprosessen vil gå lettere om den relevante terminologien er kartlagt og oversatt.

For at terminologien i basen skal bli anerkjent av de ulike fagmiljøer, er det en forutsetning at den er korrekt og i størst mulig utstrekning omfatter norske termer som allerede er i bruk i miljøene. Alle våre forslag til norske termer oversendes derfor utpekte fagmyndigheter til godkjenning. Deres gjennomgåelse av termlistene er som oftest klargjørende, men det hender naturligvis at det er flere termer i omløp for samme begrep, eller at ulike fagekspertene har divergerende mening om hva som er rett norsk term. I enkelte tilfeller kan fageksperten ha vansker med å finne en norsk term, endre oppfatning underveis eller en sjelden gang ta feil.

Status for UD-basen per oktober 1991 viser at den omfatter ca. 5 500 termer. Dette antallet skriver seg fra oversettelsen av ca. en firedel av den totale tekstmassen nevnt over. Det totale antall oppslagsformer er således for øyeblikket ca. 16 500 om man regner hovedtermer for de tre språkene. Atskillig større blir antallet om man dessuten teller med synonymer som iblant legges inn (holdbarheten av slik «synotermi» kan selvfølgelig diskuteres), kortformer, forkortelser og numerisk identitet for kjemiske stoffer.

### **Problemer i tilknytning til det terminologiske arbeidet**

Alle som har drevet med terminologisk virksomhet, vet at det ofte er problematisk å trekke grensen mellom terminologi og allmennspråk og mellom termer og kollokasjoner. Disse

problemene skal ikke berøres her. I stedet skal oppmerksomheten konsentreres om noen terminologiske problemer som er mer spesifikke for vårt arbeid.

Et overgripende problem i terminologiarbeidet, ved siden av tidspresset for å bli ferdig innen gitte frister, er at tekstmassen som skal ekserperes, er så stor og omfatter så mange og svært ulike bruksområder (emneområder). Det store antallet bruksområder gjør at man må forholde seg til mange ulike fagekspert, noe som ikke nødvendigvis er en fordel. At det ikke er mulig å spesialisere seg, vil framgå av oversikten over bruksområdene nedenfor:

- miljø
- arbeidsmiljø
- juss
- EF-rett og EF-organer
- selskapsrett
- sosiale spørsmål
- produktsikkerhet
- produktansvar
- opphavsrett
- offentlige innkjøp
- kapitalbevegelser
- finansielle tjenesteytelser
- forsikring
- tekniske handelshindringer
- konkurranse
- etableringsrett
- telekommunikasjon
- forbrukervern

Oversikten er på langt nær fullstendig, ettersom det under tekniske handelshindringer tilkommer hele 23 underområder med meget stor spennvidde, fra *kosmetikk* via *landbrukstraktorer* og *byggematerialer* til *leketøy*.

Flest problemer volder termer av følgende kategorier, selv om det gjerne er slik at flere av kategoriene opptrer samtidig i forbindelse med én og samme term:

- termer med abstrakt begrepsinnhold
- fagspråklige verb
- termer der det er begrepsulikheter mellom kildepråkene (fransk/engelsk) og mellom (ett av) disse og norsk
- «EF-termer» (EF-interne konstruksjoner uten direkte motsvar i nasjonal språkbruk).

#### Termer med abstrakt begrepsinnhold

Termen skal få anskueliggjøre de problemer som kan oppstå i forbindelse med abstrakte termer (det er bare tatt med felt som er nødvendige for å belyse problemstillingen):

F	hvd	décision
	kon	lorsqu'une "décision" a été prise

E	hvd	decision
	kon	when a "decision" has been taken
N	hvd	vedtak
	kon	når "vedtak" er gjort
.	brk	EØS
.	brk	EF-RETT
=		

(Forklaring av forkortelser: hvd = hovedterm, kon = kontekst, brk = bruksområde)

Normalt byr ikke denne termen på vansker. De fleste ville vel mene at *vedtak* eller *beslutning* er gode norske ekvivalenter. Problemet er imidlertid at termen i dette tilfellet representerer flere ulike delprosesser innen EFs institusjonsstruktur, og at det er vanskelig å skjelne realiteten bak hver av dem. I den norske oversettelsen av Roma-traktaten (1972) er det med støtte i dansk og tysk praksis, og delvis i norsk språkbruk, brukt tre ulike termer for den franske/engelske. Det skal således hete henholdsvis *vedtak* og *beslutning* når Rådet bestemmer noe, avhengig av hvor i beslutningsprosessene dette gjøres, mens domstoler tar *avgjørelser*.

Videre blir problemet om det i dette tilfellet skal registreres bare én term på fransk og engelsk. Hva gjør man i så fall med de norske ekvivalentene? Det vil neppe være riktig å føre opp én av de norske ekvivalentene som hovedterm og de andre to som synonyme termer, og uansett ville valget av norsk hovedterm måtte bli vilkårlig.

Den riktige løsningen er sannsynligvis å gå ut fra at ordet på fransk/engelsk dekker tre ulike begreper og derfor bør registreres som tre ulike termer: «*décision du Conseil/Council decision*» (no. *rådsvedtak*), «*décision de la Commission/Commission decision*» (no. *kommisjonsbeslutning*), «*décision du tribunal/court decision*» (no. *domstolsavgjørelse*).

### Fagspråklige verb

En liknende problematikk knytter seg til det fagspråklige verbet «*consulter/consult*»:

F	hvd	consulter
E	hvd	consult
N	hvd	konsultere
.	brk	EØS
.	brk	JUR
.	kom	Se kommentar under "consulter" = "høre".
=		
F	hvd	consulter
E	hvd	consult
N	hvd	rådspørre
.	brk	EØS
.	brk	JUR
.	brk	EF-RETT
.	kom	Komiteer "rådspørres".
.	kom	Se kommentar under "consulter" = "høre".
=		
F	hvd	consulter
	kon	les états membres désignent les autorités à "consulter"
E	hvd	consult
	kon	the Member States shall designate the authorities to be "consulted"
N	hvd	høre
	kon	medlemsstatene skal utpeke de myndigheter som skal "høres"

brk	EØS
brk	JUR
brk	EF-RETT
kom	"Consulter" kan også oversettes med 1) "rådspørre" (rådspørre komiteer) og 2) "konsultere" (konsultere andre, f.eks. partene i arbeidslivet). Se egne poster "rådspørre" og "konsultere".
=	

(Forklaring av forkortelser: kom = kommentar beregnet på eksterne brukere)

Skal man betrakte dette som én term eller som tre? Mye taler for det sistnevnte alternativet, selv om både fransk og engelsk bare har én betegnelse. Igjen er det begrepsmessige årsaker som veier tyngst ved avgjørelsen. Det er nemlig tre klart ulike prosesser og forskjellige aktører som skjuler seg bak termene. Ulikheten kommer godt fram i de norske ekvivalentene, så videre utdyping synes unødig.

Løsningen ovenfor avføder et nytt problem, nemlig hvilken form disse termpostene bør ha i basen. Utdragene over viser en løsning som neppe er den beste. Sannsynligvis burde hovedtermene inneholde verbet pluss et substantiv som angir målet for handlingen, f.eks. «consulte une autorité/consult an authority/høre en myndighet». Dette ville overflødiggjøre kontekstfeltet. Eventuelt burde man også i et kommentarfelt angi hvem som utfører (verbal)handlingen.

#### Termer der det er begrepsulikhet mellom kildespråkene (fransk/engelsk) og mellom (ett av) disse og norsk

Iblant forekommer det i tekstene begrepsulikhet mellom de to kildespråkene og/eller mellom (ett av) dem og norsk. Eksemplet under viser begrepsulikhet mellom fransk på den ene siden og engelsk og norsk på den annen:

F	hvd	sanction
E	hvd	penalty
N	hvd	straffetiltak
N	syn	straffebestemmelse
.	brk	EØS
.	brk	JUR
=		
F	hvd	sanction
E	hvd	sanction
N	hvd	sanksjon
.	brk	EØS
.	brk	JUR
=		

(Forklaring av forkortelser: syn = synonym)

Som utdragene viser, har fransk «sanction» en videre ekstensjon (begrepsomfang) enn engelsk «sanction» og rommer både negative og positive tiltak. På engelsk, som på norsk, er i våre tekster det tilsvarende begrepsområdet oppdelt på to termer, den ene for negative tiltak («penalty»), den andre for positive («sanction»).

Et liknende tilfelle er vist under, men denne gang er det norsk som skiller seg fra fransk og engelsk:

F	hvd	accident de travail
E	hvd	accident at work
E	syn	industrial accident
N	hvd	arbeidsulykke
.	brk	EØS
.	brk	ARBEIDSMILJØ
.	brk	FORSIKRING
.	mrk	Se "accident de travail" = "yrkesskade" innenfor SOSIAL.
=		
F	hvd	accident de travail
E	hvd	accident at work
N	hvd	yrkesskade
.	brk	EØS
.	brk	SOSIAL
.	mrk	Se "accident de travail" = "arbeidsulykke" innenfor FORSIKRING og ARBEIDSMILJØ.
=		

(Forklaring av forkortelser: mrk = kommentar beregnet på Oversettelsesenshetens brukere)

I henhold til fagmiljøene skilles det i norsk lovgivning mellom «arbeidsulykke» innen fagområdet *arbeidsmiljø* og *forsikring* på den ene siden og «yrkesskade» innenfor *sosiallovgivning* på den annen. De to termene fokuserer på ulike aspekter av samme hendelse, henholdsvis årsak og selve hendelsen (*arbeidsulykke*) og resultat (*yrkesskade*). Sannsynligvis har de fleste med norsk som morsmål også en følelse av at den førstnevnte kan, men ikke nødvendigvis må, føre til den sistnevnte. På fransk (og engelsk?) gjøres tydeligvis ikke denne distinksjonen i lovspråket, selv om den naturligvis kan uttrykkes i allmennspråket. I franske og engelske ordbøker finner man da også betydningsvarianten «skade» under oppslagsordet «accident».

#### «EF-termer»

Med «EF-term» menes her en term som har oppstått innenfor Fellesskapets institusjoner, f. eks. innenfor lovgivning eller i forbindelse med iverksettelse av lover. Man kunne med like stor rett kalle dem «direktiv-termer», ettersom de nok først og fremst er et utslag av direktivenes indirekte rettskraft. Ettersom direktivene jo skal innarbeides i de ulike medlemsstaters nasjonale lovgivning, er det ofte viktig for EFs lovgivere å mynte termer som fanger opp praksis og/eller nomenklatur i alle medlemsstatene. Slike termer må rimeligvis få svært vid ekstensjon og blir dermed generelle, noe som altså er tilsiktet. Av jurister oppfattes en slik uttrykksform muligens som fordelaktig, men for terminologen (og for folk flest) er den et mareritt, ettersom det blir tilnærmet umulig å avgrense termens ekstensjon. Følgende eksempel burde vise dette:

F	hvd	diplôme
E	hvd	diploma
N	hvd	diplom
.	brk	EØS
.	brk	SELSKAPSRETT
.	brk	ETABLERINGSRETT
.	mrk	N hvd endret fra "vitnemål" til "diplom" av Lund i KUF.
=		

På fransk og engelsk kan man uttrykke seg svært generelt ved hjelp av den oppførte termen. Tilsynelatende er norsk «vitnemål» ekvivalent når det gjelder utdanning og opplæring innen offentlige eller private institusjoner, ettersom det er en overterm for ulike typer eksamensbevis o.l. Problemet er imidlertid at «vitnemål» ikke omfatter kursbevis og opplæringsdiplomer for utdanning *uten eksamen eller prøve*, noe «diplôme/diploma» gjør på fransk/engelsk. «Vitnemål» har ut fra denne betraktningen en snevrere ekstensjon enn kildespråkernes term. Vår fagkontakt («KUF:Lund») har i dette tilfellet valgt en pragmatisk løsning; det likelydende «diplom» blir norsk ekvivalent i EF-sammenheng.

Den valgte løsningen er uheldig ettersom den ikke tar hensyn til det tilvante begrepsinnholdet av «diplom» på norsk. Diplomer forbindes med skirenn eller uforpliktende kurs av noen få dagers varighet. I direktivet forekommer det imidlertid en oppramsende definisjon av termen, og når direktivet eventuelt skal innarbeides i norsk lov, vil sannsynligvis de tradisjonelle norske termene erstatte EF-termen.

Til slutt vises et eksempel på en EF-term som ikke har norsk ekvivalent:

F	hvd	attestation d'adquation
E	hvd	certificate of adequacy
N	hvd	yyy
N	syn	tjenlighetssertifikat
.	brk	EØS
.	brk	THH
.	mrk	Strand i DAT kjenner ikke noe norsk sertifikat med tilsvarende betegnelse. Han mener at den norske termen over ev. kan innføres. Han anbefaler ellers omskriving som i den danske versjonen.

=

Termen over brukes om et sertifikat som innenfor EFs indre marked skal bekrefte at et produkt tilfredsstillende visse spesifiserte krav. Som det framgår av merknadsfeltet, har vi i Norge i dag ikke noe slikt sertifikat. Den norske termen er konstruert under oversettelsesarbeidet, og ettersom det foreløpig er uvisst om den vil bli benyttet i framtida, står feltet «N hvd» tomt inntil videre. Det er så langt også uvisst nøyaktig hvilket begrepsområde termen dekker.

Det er først og fremst når det gjelder det som her er kalt «EF-termer», at oversettelsen av avtaletekstene kan komme til å påvirke norsk terminologi og språk. Stort sett bare når termene på fransk og engelsk er tilsiktet generelle eller dekker et begrep som utelukkende eksisterer innenfor EFs institusjoner, vil det i denne sammenheng være behov for å skape nye norske termer. Det er i slike tilfeller, liksom ved alt terminologisk arbeid, viktig at oversettere, terminologer og fageksperter anstrender seg for å finne den norske ekvivalenten som har (størst) oppslutning i fagmiljøene, og at de bestreber seg på å finne gode *norske* avløsertermer. Ellers kan det gå som det gjorde da fagmiljøet ønsket ut den norske ekvivalenten til termen «organisme notifié/notified body»: «notifisert organ». Termen er både ugjennomsiktig og syntaktisk tvetydig. Uten å gå nærmere inn på det bakenforliggende begrepsinnhold må det være lov å spørre: har noen «notifisert» organet om noe, eller er noen «notifisert» om organet? I dette tilfellet har vi valgt å utfordre fagmiljøet, og vårt forslag til norsk term er «meldt organ». På grunn av sitt norske språkstoff er den gjennomsiktig, og dessuten entydig. Likeledes finner den sin naturlige plass blant «melde» («notifier/notify»), «melding» («notification/notification») og «melder» («notifiant/notifier»).

Oversettelsen av avtaletekstene er et stort prosjekt, og arbeidet utføres under et visst tidspress. Disse rammene gir ikke tid til å dvele for lenge ved enkeltheter. I det terminologis-

ke arbeidet er vi dessuten for det meste bundet av den terminologi som allerede er fasttømret i de ulike fagmiljøer. I enkelte tilfeller har vi imidlertid mulighet til å påvirke ny eller eksisterende terminologi. Som de siste eksemplene viser, etterstreber vi da gjennomsiktighet og entydighet, samt bruk av norsk språkstoff og enhetlig behandling av beslektede termer. På denne måten ønsker vi å gi oversettelsene en høy kvalitet og et norsk særpreg.



## Stikkordliste

Stikkordlisten er ment som en hjelp for leserne til å finne fram til hvor forskjellige emner er omtalt i rapporten. Stikkordene er ordnet alfabetisk etter norsk alfabetrekkefølge. Ordene er ført opp i den formen de er brukt i artiklene. Det betyr at både dansk, norsk og svensk forekommer i lista, f eks vil en finne både begreb, begrep og begrepp. Forfatterens stavemåte er stort sett beholdt uendret, og en kan dermed finne f eks både tesaurus og thesaurus. Ordningen er gjort dels av enkelhets skyld og for å unngå ekvivalens- og normeringsproblemer. Det er lagt vekt på å få med frekvente forkortinger/initialord i lista.

- afsender 269
- afsenderholdning 269
- agglutinerende språk 437
- aktiv användning 347
- aktiva ordböcker 16
- aktiv bruk 90
- aktive ordbøger 483
- aktivt speciallexikon 347
- albog 275
- alfabetisering 184
- alfabetisk-semasiologisk 341
- alfanumerisk sortering 188
- allmennspråk 211
- allmänna alfabetiska ordböcker 14, 23
- allusjon 84
- almindelig bog 267
- alordbøger 275, 522
- anakronisme 211
- analogikriterie 141
- "anbeheitelse"-ord 56
- anisomorfi 480
- antonymordböcker 14
- appellativ 191, 372
- appellativisert proprium 191
- areala drag 156
- aristotelisk definition 146
- avgränsningsprincip 359
- avledning 467
- avledningsmorfem 34
- avleiing 320
- avstavningsordböcker 14
- avveiningsmarkør 217
- baklängesordböcker 14
- bas 353
- basisord 289
- begreb 524
- begrebsdefinitionsordbøger 264
- begrebsomfang 145
- begrep 81
- begreppsklasse 167
- begreppsordböcker 166
- begrepsanalytisk metode 503
- begrepskjennetegn 222
- begrepsulikhet 545
- begränsede kollokationer 289
- belägg 19
- betoning 467
- betydelsebeskrivning 18
- betydelsefält 166
- betydelseglidning 518
- betydelsemoment 138
- betydelseutveckling 232
- betydningslücke 214
- betydningslån 308
- betydningsnummer 82, 215
- betydningsomfang 216
- bifunktionella ordböcker 17
- bildordböcker 14
- bilingvale ordbøger 272
- billedlig 84
- bitone 471
- bokmål 53
- bokstavsordning 184
- brugervejledning 271
- brugssituation 270
- bruklighetsmarkering 19

- bruksinformation 466  
bruksområde 543  
CAT (computer-aided translation) 348  
central betydelse 351  
cirkelhenvisninger 529  
citatlån 361  
citatordböcker 15  
COBUILD-format 349  
componential analysis 503  
databehandlingsterminologi 409  
definiendum 146  
definiens 146  
definition 15, 144, 349  
definitionsordböcker 264  
delbetydelse 20  
deleksikalisering 291  
delfagsordböcker 281  
delvis ekvivalens 502  
demonstration 19  
denotasjon 82, 191  
deskription 46  
deskriptiv 43, 515  
deskriptiva ordböcker 13  
diakritiska tecken 188  
diakron 180  
diakroniska ordböcker 11  
dialektform 340  
dialektsleksikografi 339  
dialektordböcker 12  
diasituative kriterier 180  
diastratiske kriterier 180  
diatekniske forkortelser 246  
diatekniske markeringer 525  
diatopiske kriterier 180  
dictionary 81  
Dictionary Entry Parser (DEP) 428  
Dictionary Access Method (DAM) 426  
differentia 146  
diftongering 178  
diminutiv 178  
direkte ekvivalens 502  
direktlån 361  
DK87-90 492  
dobbelte lån 310  
dobbellematisering 181  
egennamn 33  
egenskapsnamn 467  
ekstension 541, 545  
ekstensionell definition 146  
ekte term 213  
ekvivalent 15, 435  
emneområde 543  
emotive vulgarismer 434  
encyclopaedi 81  
encyklopediska ordböcker 10  
encyklopedologi 219  
encyklopædiske ordbøger 264, 276  
enkeltfagsordböcker 281  
enspråkiga ordböcker 15  
etymologi 19, 232  
etymologiske opplysninger 345  
etymologiske tyding 191  
etymologisk fagordbog 265  
EUROTRA 487  
evfemisme 434  
excerpering 370  
excerpt 369  
execerption 180  
exegetiska ordböcker 11  
exempel 18, 159  
explanatory combinatorial dictionaries 353  
explikation 18  
fackordböcker 13  
fagleksikografi 248  
fagord 244  
fagordböcker 262  
fagspråklige verb 544  
fagsprog 524  
fagterm 254  
familielighed 150  
familiær 432  
faste ordforbindelser 94  
fast sammensetning 77  
fast uttrykk 83  
feltkode 425  
feltteori 81  
fennicisme 516  
fileringsmetoder 186  
finalalfabetisk sortering 188  
finlandismer 516  
flerfagsordböcker 281  
flerordsenhed 484  
flerspråkiga ordböcker 17  
folkelig 433  
fonetiska former 188  
forkortelse 452

- forkortelsesordbøger 265  
 forkortelsespunktum 453  
 form 310  
 fornsvenska 231  
 forord 271  
 frase 80, 360  
 frasem 83  
 fraseoleksem 485  
 fraseologi 83, 467, 476  
 fraseologiske udtryk 484  
 frekvensordbøger 14  
 fremmedordbøger 224  
 funktion 524  
 fællesord 256  
 förtäckt normativa ordböcker 13  
 förtäckt normativitet 47  
 gammalnorsk 315  
 gammeldansk 455  
 gardsfolknamn 194  
 generisk 450  
 genetiska drag 156  
 genomskinlighet 35  
 genusbegreb 147  
 genus proximum 222  
 gerundium 534  
 glosor 167  
 glossar 316  
 grafie 178  
 graford 33  
 grammatiska kollokationer 354, 476  
 "grove" begreper 434  
 grunnform 310  
 grundmorfem 34  
 grunnord 90  
 hapax legomenon 343, 383  
 henvisningslemma 181  
 henvisningsmarkør 297  
 henvisningsord 301  
 heterografe homofoner 176  
 heteronym 341  
 historiska ordböcker 11, 231  
 historiska samtidsordböcker 12  
 historiske ordbøger 390  
 historiske ordbøker 231  
 historisk leksikografi 177  
 homograf 310, 442, 453, 468  
 homografdeling 180  
 homografiske former 396  
 homografnummer 451  
 homografordbøcker 14  
 homonymordbøger 176  
 homonymdifferentiering 176  
 homonymer 81  
 homonymi 20  
 hovedfagsordbøger 281  
 hovedlemma 181  
 hovedopslagsord 298  
 hovudform 53  
 husfolknamn 194  
 huvud 475  
 huvudord 232  
 hyperonym 14  
 hyponym 14  
 hårda lexikala funktioner 354  
 ideel fonemfølge 179  
 ideel grafemfølge 179  
 ideelt grafemsystem 179  
 ideologiske ordböcker 166  
 idiom 18, 84, 442  
 idiomatiske udtryk 485  
 idiomordbøcker 15  
 idiosynkratiske kombinationer 288  
 idiotikonprinsippet 342  
 ikke-fagord 251  
 index 11  
 indholdstræk 290  
 indirekte lån 308  
 indre selektion 268  
 informant 381  
 inkolentnamn 194  
 inkompatibilitet 147  
 inlæringslexikon 347  
 innbyggjarnemne 192  
 inngang 224  
 innskrivingsformat 321  
 intensjon 541  
 intensionell definition 146  
 internationalisme 255  
 inverterad följd 186  
 jamning 341  
 jysk 331  
 kanoniska normalformer 188  
 kategorimedlemskab 150  
 kildefil 424  
 kildehenvisning 311  
 kjeldetilvising 344

- klammeform 53  
 klassifikasjon af ordbøger 263  
 kodeliste 79  
 kohyponym 14, 303  
 kollokasjon 83, 344, 435, 442  
 kollokation 18, 280, 288, 353, 489  
 kollokationsordböcker 15  
 kollokator 353  
 kombinasjoner 84  
 kombinationsordböcker 353  
 komposisjonsfuge 341  
 komposita 248  
 konkordans 11, 370  
 konnotasjon 191, 432  
 konnotationsfri kommunikation 25  
 konsonantstam 437  
 konstruksjoner 18, 84  
 konstruksjonsordböcker 15  
 konstruktlemma 179  
 kontekst 371  
 kontrastiv analyse 156  
 konversasjonsleksikon 81  
 korpus 11  
 krydshenvisning 280, 528  
 KWIC (Key Word in Context) 492  
 källspråk 16, 437  
 kärnverb 157  
 ledevarietet 177  
 legaldefinisjon 212  
 leksikalske funksjoner 288  
 leksikalske vulgarismer 435  
 leksikografi 76  
 leksikogrammatiske restriksjoner 289  
 leksikon 488  
 lemma 268  
 lemmeegn 278  
 lematisering 33, 177  
 lexem 138, 348  
 lexikala funksjoner 353  
 lexikala kollokasjoner 354  
 Lexical Data Base (LDB) 426  
 lexical functions 353  
 lexikal profil 157  
 Lexical Query Language 423  
 leksikaliserad enhet 22  
 leksikalisering 35  
 leksikaliseringshierarki 157  
 leksikaliska enheter 34  
 leksikografi 31  
 leksikologi 32  
 leksikon 31  
 leksikonkomponent 30  
 ljudlängd 467  
 logogram 188  
 lukkede enheder 493, 494  
 læreboknormal 53  
 löpande ord 370  
 løs sammensetning 77  
 låneord 170  
 lånord 520  
 Machine Translation (MT) 488  
 makrostruktur 19, 267, 528  
 maksimale oversettelser 536  
 maksimerende fagsproglige ordbøger 280  
 markerade drag 156  
 materielle markører 434  
 Meaning-Text-theory 353  
 mediopassiv 98  
 megastruktur 267  
 mellomnorsk 315  
 meningskompositionalitet 146  
 metafor 84  
 metaleksikografisk forskning 177  
 metasproglig funksjon 264  
 mikrostruktur 19, 267, 475, 528  
 mindre ord 386  
 minimerende fagsproglige ordbøger 280  
 mjuka lexikala funksjoner 354  
 modellord 520  
 modifierade direktlån 362  
 modtager 270  
 monoftongering 178  
 monofunksjonella ordböcker 17  
 morfofonemiske vulgarismer 434  
 MT (Machine Translation) 488  
 multilingvalitet 488  
 målform 53  
 målspråk 16  
 namnesemantisk tyding 191  
 namnprefix 187  
 nasjonalordbøker 235  
 Natural Language Processing (NLP) 488  
 neosemantisering 213  
 nesting 23  
 nicheleksikografer 481  
 niching 23

- NLP (Natural Language Processing) 488  
 nollsteg 476  
 nomen communis 194  
 norm 432, 515  
 normativa ordböcker 13, 515  
 normativ funksjon 50  
 normativ intention 50  
 normativitet 223  
 normbundethed 486  
 normering 35, 237, 341  
 norvagisme 316  
 novasjon 432  
 nukleära verb 161  
 ny betydelse 233  
 nybildning 362  
 nynorsk 53  
 nyord 79  
 nyordbok 12, 308  
 nytt språk 232  
 nyutvikling 317  
 nøgen henvisning 387  
 nøsting 341  
 nøytrale betegnelser 435  
 obskøne ord 434  
 omarkerade drag 156  
 omskrivning 84  
 onomasiologiska ordböcker 14  
 onomastic material 202  
 opplysningstype 223  
 oppslagsform 340  
 oppslagsord 82  
 ord 81  
 ordartikkel 82  
 ordbehandlingsprogram 348  
 ordbog 261  
 ordbogsartikkel 451  
 ordbogsgrammatik 280  
 ordbok 11, 54, 81  
 ordbokssedler 403  
 ordbokstypologi 24  
 ordfogning 466  
 ord-før-ord-princip 188  
 ordgrupper 360  
 ordliste 54, 81  
 ordningsform 187  
 ordningsmåte 223  
 ordspråk 471  
 ordspråksordbøcker 15  
 overført betydning 84  
 oversettelsesekvivalent 502, 533  
 oversettingslån 308  
 oversættelsesproblem 483  
 pakning 299  
 paradigmatiske informasjon i ordbøcker 18  
 parafras 15  
 parafrase 536  
 paronymi 179  
 passiv anvendning 347  
 passiv ordbok 16, 308, 319  
 plusord 249  
 polylingvale ordbøger 272  
 polysemi 20, 81, 137, 301, 5 25  
 polysemimønster 158  
 preskription 46  
 preskriptiv 43  
 produktivitetskriterie 141  
 proprier 190, 221, 277, 319  
 propriumsrelaterte appellativ 192  
 prototype 149  
 prototypikalitet 159  
 proveniens 460  
 psevdolån 308  
 rammeparadigme 310  
 rangordsfølj 187  
 realia 221  
 realopplysninger 220  
 redaktionsexempel 19  
 redealfabetisk 253  
 redigeringsprogram 474  
 referensologi 226  
 reguljær polysemi 140  
 rekommendation 47  
 relationsforskydning 301  
 relevantsteoretisk tilnærming 503  
 reservation 47  
 restringerte kollokasjoner 80  
 rettskriving 53  
 rettskrivingsordliste 53  
 riktighetskala 435  
 rimordbøcker 14  
 rækkesammensætning 325  
 sagordbøger 264, 522  
 samansetning 320  
 samleartikler 298  
 sammansatt ord 466  
 sammansættingsførl 232

- sammensetning 533  
 samtidsordböcker 11  
 samtidsspråket 231  
 sandhi 433  
 satsinledande block 35  
 seddelsamlinger 390  
 selektionsregler 489  
 selektiv ordbog 478  
 semantisk differenteriering 158  
 semantisk-syntaktiska restriktioner 352  
 semantiska fält 157  
 semantisk valens 98  
 semasiologiska ordböcker 14  
 semem 138  
 sememisk forbindelighed 489  
 semifagord 251  
 semilattice 147  
 semiologisk nivå 223  
 semiologisk trekant 219  
 semisk forbindelighed 489  
 setningsomgivelser 94  
 SGML 224, 447  
 SGML-format 457  
 sideform 53  
 sidelemmal 181  
 sigel 451  
 simplex-verber 490  
 simplexord 35  
 sirkeldefinisjon 77  
 skjellsordbøker 345  
 skriftnormal 235  
 slang 519  
 slangordböcker 13  
 smygnormativa ordböcker 13  
 søkeverktøy 407  
 sorteringsnyckel 188  
 sprogordbøger 276, 336  
 språklige tecken 14  
 språkprov 369  
 språkriktighet 212  
 språkspecifika drag 156  
 stadieväxling 437  
 stammestrek 427  
 standardisering 509  
 standard østnorsk 433  
 stavekontroll 65  
 stavningsordböcker 14  
 stavningsprinciper 188  
 stednavne 455  
 stereotypi 144, 149  
 stikkord 223  
 stikkordutvalg 223  
 stilmarkører 303  
 strukturell typologi 267  
 stykke 169  
 style checkers 348  
 større ord 386  
 støtteverbumbkonstruksjoner 489  
 stående talesätt 471  
 sublemmaer 485  
 substitutionsprinsippet 77  
 support verb constructions (SVC) 489  
 suprasegmentale former 434  
 svake vulgarismer 435  
 svarabhaktivokal 319  
 SVC (support verb constructions) 489  
 synkroniska ordböcker 11  
 synkront-semasiologiske ordbøger 177  
 synonym 14, 145, 296  
 synonymgruppe 297  
 synonymi 145  
 synonymordböcker 15  
 synotermer 542  
 syntagmatiska ordböcker 14  
 syntagmatisk information i ordböcker 18  
 syntaktisk valens 98  
 syntetiskt språk 437  
 systematisk-onomasiologisk ordning 341  
 särnusvenska ord 233  
 sætnings- og billedordbøger 265  
 tabu 434  
 tabuskala 435  
 talemåte 83  
 talesprog 381  
 talesprogsamling 391  
 tecken-för-tecken-princip 188  
 tegninger 527  
 tekststreng 407  
 term 211  
 termekvivalent 361  
 terminografisk 33  
 terminologi 24, 211, 214, 509  
 terminologiordbøger 266  
 terminologisering 212  
 terminologiska ordlistor 24  
 terminologisk system 33

- tertium comparationis 178  
thesaurus 14, 144  
textform 186  
thesaurus 265  
tilde 427  
tolkordlistor 357  
tomme verber 290  
tonem 59, 82  
topografisk listning 335  
tosprogede fagordbøger 260  
tosprogede oversættelsesordbøger 260  
tospråklige ordbøger 435  
transferordbøger 488  
trykkaksent 427  
trykkmarkering 59  
tvetydighet 511  
tvärspråklig lexikologi 156  
tvåspråkiga ordböcker 466  
tyd 451  
tydingsdifferensiering 237  
typografisk kodning 450  
typologi 10  
typologiska drag 156  
UDK 261  
universella drag 156  
uppslagsform 186  
utbytesprov 139  
uttale 341  
uttalsordböcker 14  
uttryck 81  
utvidelse 84  
valensbetingede flerordsenheder 487  
valensteori 81  
valensuppgifter 351  
varianttecken 188  
variantöversättning 475  
vokalharmoni 319  
vokalstam 437  
vulgarisme 432  
vulgært 433  
vurderingsskala 435  
WordSmith 423  
ydre selektion 268  
zeugma 140  
ækvivalensangivelse 254  
ækvivalensproblem 124  
ækivalenttegn 279  
öppet normativa ordböcker 13  
översättningsdubletter 362  
översättningslån 361  
åbne enheder 493, 494  
åbne kollokationer 289  
ålderdomligt språk 232

## Refererte ordbøker

Ordbøkene står oppført i den form de faktisk er brukt i artiklene. Skrivemåte og forkorting må dermed tilskrives den enkelte forfatter. Det medfører at samme ordbok kan være oppført på forskjellige steder.

- Advanced Learner's Dictionary 18
- Aschehoug og Gyldendals Store Norske Leksikon 221
- BBI Combinatory Dictionary 94
- Black's Law Dictionary 251
- Bokmålsordboka 61, 65, 83, 151, 213, 220, 240, 309, 407, 409, 423, 433
- Bonniers svenska ordbok (BSO, 1986) 43
- Cobuild-ordboka 88
- Collins Cobuild English Language Dictionary 297, 304
- Collins Concise Dictionary and Thesaurus 297, 298
- Collins Thesaurus 409
- Collins' tvåspråkiga ordböcker 17
- Cyclopaedia, or Universal Dictionary of Arts and Sciences 219
- Danmarks Stednavne 456
- Dansk-engelsk ordbog 17
- Dansk Sprogbrug 94
- Dansk synonymordbog 296, 298, 300
- Den arnamagnæanske kommissions ordbog 385
- Dictionary of Computing 409
- Duden Deutsches Universalwörterbuch 513
- Duden Sinn- und sachverwandte Wörter 296, 300, 301, 302, 303
- Duden Bedeutungswörterbuch 297
- Encyclopædia Britannica 219
- Engelsk-dansk ordbog 17, 409
- Engelsk-norsk ordbok 409
- Engelsk-norsk ordbok 435
- Esseltes svensk-engelsk ordbok 442
- Fremmedordboka 409, 423
- Fritzners ordbok 316
- Frühneuhochdeutsches Wörterbuch 177
- Frühneuhochdeutsches Glossar 177
- Gamalnorsk ordbok 196
- Gammeldansk Ordbog 196, 455
- Grimm, Deutsches Wörterbuch 231
- Gulbrandsen: Juridisk leksikon 213
- IBM Norge: Synonymordbok 409
- Illustrerad svensk ordbok 138, 232, 466
- Íslensk orðabók 91
- Íslensk-dönsk orðabók 91



- Juridisk ordbog, dansk-tysk 118  
 Jysk Ordbog 237, 325, 343  
 Kunnskapsforlaget Norsk-fransk ordbok 534  
 Lademanns Leksikon 221  
 LEXIN-ordböckerna 18  
 Longman Dictionary of Contemporary English 18, 138  
 Longman Dictionary of the English Language 296  
 Matias Skard: Nynorsk ordbok 61  
 McGraw-Hill Dictionary of Scientific and Technical Terms 514  
 Middelnederlandsch Woordenboek (Middle Dutch Dictionary) 202  
 New English Dictionary 231, 370  
 Norrøn ordbok 196  
 Norsk Dataordbok 221  
 Norsk Ordbok 190, 194, 195, 235, 316, 339, 342, 406  
 Norsk landbruksordbok 240  
 Norsk Ordbog (Ross) 195  
 Norsk riksmålsordbok 83  
 Norsk Ordbog (Aasen) 3, 195  
 Norsk-engelsk ordbok 409, 423  
 Norstedts Stora engelsk-svenska ordboken 17  
 Nudansk Ordbog 146, 195, 326  
 Nusvensk ordbok 369, 465  
 Nykysuomen sanakirja 465, 466, 468  
 Nynorsk etymologisk ordbok 196  
 Nynorskordboka 61, 196, 240, 342, 407  
 Ord för ord 169, 296, 298, 300, 302, 303  
 Orðabók Háskólans 92, 370  
 Ordbog over det gamle norske Sprog 196  
 Ordbog over det danske sprog 195, 232, 297, 325  
 Ordbog til det ældre danske sprog 195  
 Ordbog over det danske Sprog 151  
 Ordlista för tolkar 357  
 Oxford English Dictionary 12, 405  
 Oxford Advanced Learner's Dictionary 512, 513  
 På godt norsk 409  
 Politikens Nudansk Ordbog 297, 300, 301  
 Politikens Synonymordbog 296, 297, 298, 299, 301, 302  
 Retskrivningsordbogen i Danmark 61  
 Riksmålsordboken 433  
 Roget's Thesaurus 15, 409  
 Ross 316  
 Schweizerisches Idiotikon 235  
 Scottish National Dictionary 235  
 Special-ordbog inden for revision, regnskabsvæsen m.v. 118  
 Stora svensk-engelska ordboken 17, 474, 475  
 Store norske ordbok 213  
 Suomen kielen perussanakirja 468, 516  
 Svensk-Engelsk Fackordbock 252

- Svensk ordbok 43, 137, 146, 232, 297, 304, 347, 467, 468  
Svensk etymologisk ordbok 196, 378  
Svensk handordbok 94, 353, 466  
Svenska Akademiens ordbok 12, 195, 231, 369  
Svenska Akademiens ordlista 13, 43, 138, 466, 468  
Svenska ord 351  
Svenskt-finskt Lexikon 465  
Tanums store rettskrivningordbok 65, 221  
TFSNOR (engelsk-norsk dataterminologi) 409  
The American College Encyclopedic Dictionary 220  
The BBI Combinatory Dictionary of English 293, 347  
The Concise Oxford Dictionary 513  
The Penguin Modern Guide to Synonyms and Related Words 296  
The Oxford Thesaurus 296  
Thesaurus Linguae Latinae 370  
Thüringerordboka 342  
Trésor de la langue française 141  
van Dales nederländsk-tyska och tysk-nederländska ordböcker 17  
van Dales svensk-nederlandske ordbok 441  
Vroegmiddelnederlands Woordenboek (Dictionary of Early Middle Dutch 200  
Webster 144, 409  
Webster's New Twentieth-Century Dictionary 514  
Webster's Collegiate Thesaurus 296, 297, 298, 302, 303  
Webster's Dictionary of Synonyms 296  
Webster's New Twentieth-Century Dictionary 512  
Websters ordbok 139  
Woordenboek der Nederlandsche Taal 202, 231

## Oversikt over forfatterne

- Akø, Jørn-Otto, prosjektassistent, f. 1958, Universitetet i Oslo, Avdeling for leksikografi, Boks 1001 Blindern, N-0315 Oslo
- Ásta Svavarsdóttir, ordboksredaktør, f. 1955, Orðabók Háskólans, Neshaga 16, IS-107 Reykjavík
- Baustad, Jostein, student ved Norges Tekniske Høgskole i Trondheim, linje for datateknikk og telematikk, f. 1965, Havnabakken 9, N-0874 Oslo
- Becker-Christensen, Christian, ordbogsredaktør, f. 1944, Politikens Forlag A/S, Vestergade 26, DK-1456 København K
- Bergenholtz, Henning, professor, f. 1944, Handelshøjskolen i Århus, Fuglesangs Allé 4, DK-8210 Århus V
- Berkov, Valerij, f. 1929, Universitetet i Leningrad, B. Okhtinskij pr 6-196, 195027 Leningrad
- Boje, Frede, cand.mag., f. 1937, Handelshøjskolen i København, Institut for tysk, Dalgas Have 15, DK-2000 Frederiksberg
- Braasch, Anna, cand.mag., f. 1946, Center for Sprogteknologi, Njalsgade 80, DK-2300 København S
- Chaffey, Patrick Nigel, universitetslektor, f. 1938, Universitetet i Oslo, Institutt for britiske og amerikanske studier, Boks 1003 Blindern, N-0315 Oslo
- Degnbol, Helle, ordbogsredaktør, f. 1945, Ordbog over det norrøne prosasprog, Njalsgade 76, DK-2300 København S
- de Groot, Hans, forsker, f. 1959, Skandinaviske Institutionen, Universitetet i Groningen, Postbus 716, NL-9700 AS Groningen
- Duvå, Grete, f. 1942, Handelshøjskolen i Århus, Strandvænget 6, DK-8240 Risskov
- Ellingsve, Eli, førsteteknisk, f. 1952, Norsk termbank, Strømg. 53, N-5007 Bergen
- Engh, Jan, f. 1946, Ullevålsv. 56 C, N-0454 Oslo
- Fjeld, Ruth Vatvedt, stipendiat, f. 1948, Universitetet i Oslo, Avdeling for leksikografi, Boks 1001 Blindern, N-0315 Oslo
- Frøili, Jorunn, førsteteknisk, f. 1955, Universitetet i Oslo, Institutt for lingvistik og filosofi, Avdeling for tolkeutdanning, Postboks 1102 Blindern, N-0317 Oslo
- Gavare, Rolf, docent, f. 1945, Göteborgs universitet, Institutionen för svenska språket, S-41298 Göteborg
- Gellestam, Martin, docent, f. 1936, Göteborgs universitet, Institutionen för svenska språket, S-41298 Göteborg
- Graedler, Anne-Line, utdanningsstipendiat (NAVF), f. 1960, Universitetet i Oslo, Institutt for britiske og amerikanske studier, Boks 1003 Blindern, N-0315 Oslo
- Guðrún Ása Grímsdóttir, stud. mag., f. 1948, Ordbog over det norrøne prosasprog, Njalsgade 76, DK-2300 København S
- Guðrún Kvaran, redaktør, Orðabók Háskólans, Neshaga 16, IS-107 Reykjavík
- Gundersen, Dag, professor, f. 1928, Universitetet i Oslo, Avdeling for leksikografi, Boks 1001 Blindern, N-0315 Oslo
- Hageberg, Arnbjörg, amanuensis, f. 1935, Universitetet i Oslo, Avdeling for leksikografi, Boks 1001 Blindern, N-0315 Oslo
- Helleland, Botolv, førsteamanuensis, namnekonsulent, f. 1940, Universitetet i Oslo, Avdeling for namnegranskning, Boks 1011 Blindern, N-0315 Oslo

- Henriksen, Petter, sjefredaktør, f. 1949, Hjemmets Bokforlag AS, Kr. Augusts gt. 10, N-0164 Oslo
- Hoel, Jan, førstekonsulent/tekstrevisor, f. 1951, Det Kgl. Norske Utenriksdepartement, Postboks 8114 Dep., N-0032 Oslo
- Hovdenak, Marit, konsulent, f. 1950, Norsk språkråd, Postboks 8IS-107 Dep, N-0032 Oslo
- Jacobsen, Bent Chr., ordbogsredaktør, f. 1944, Ordbog over det norrøne prosasprog, Njalsgade 76, DK-2300 København S
- Jacobsen, Tove, førsteamanuensis, f. 1944, Romansk institutt, Avd. A: Fransk, Universitetet i Bergen, Sydnesplass 9, N-5007 Bergen
- Jenstad, Tor Erik, sekretær for namnekonsulenttenesta i Midt-Noreg, f. 1956, Universitetet i Trondheim, Nordisk institutt, N-7055 Dragvoll
- Jón Hilmar Jónsson, ordboksredaktør, f. 1947, Orðabók Háskólans, Neshaga 16, IS-107 Reykjavík
- Jonsson, Hans, ordbokschef, professor, f. 1928, Svenska Akademiens ordbok, Box 3, S-221 00 Lund
- Jørgensen, Merete K., konst. ordbogsleder, f. 1951, Gammeldansk Ordbog, Njalsgade 80, DK-2300 København S
- Karlsson, Göran, professor emeritus, f. 1917, Tavastgatan 4 B, 38, SF-20500 Åbo
- Knudsen, Jette, ordbogsredaktør, f. 1956, Ordbog over det norrøne prosasprog, Njalsgade 76, DK-2300 København S
- Kristín Bjarnadóttir, ordboksredaktør, f. 1950, Orðabók Háskólans, Neshaga 16, IS-107 Reykjavík
- Lange, Sven, docent, f. 1943, Institutionen för nordiska språk, Stockholms universitet, S-106 91 Stockholm
- Laursen, Anna-Lise, lektor, f. 1946, Handelshøjskolen i Århus, Nørre Allé 100, 3. tv., DK-8000 Århus C
- Lenstrup, Lita, lektor, f. 1940, Handelshøjskolen i København, Dalgas Have 15, DK-2000 Frederiksberg
- Maidahl, Lisbeth, lektor, f. 1946, Handelshøjskolen i Århus, Saralystvej 2, DK-8270 Højbjerg
- Malmgren, Sven-Göran, högskolelektor, f. 1943, Göteborgs universitet, Institutionen för svenska språket, S-41298 Göteborg
- Martola, Nina, f. 1955, Forskningscentalen för de inhemska språken, Sörnäs strandväg 25, SF-00500 Helsingfors
- Mikkelsen, Hans Kristian, seniorstipendiat, f. 1955, Handelshøjskolen i Århus, Russisk institut, Fuglesangs Allé 4, DK-8210 Århus V
- Mogensen, Jens Erik, universitetsstipendiat, f. 1964, Københavns Universitet, Institut for germansk filologi, Njalsgade 80, trappe 16, 2. etage, DK-2300 København S
- Nielsen, Sandro, adjunkt, f. 1961, Handelshøjskole Syd, Østervang 2, DK-6800 Varde
- Nikula, Kristina, biträdande professor, f. 1943, Tammerfors universitet, Nordisk filologi, Box 607, SF-33101 Tammerfors
- Niska, Helge, byrådirektör, fil.kand., f. 1945, Stockholms Universitet, Tol- och översättarinstitutet, S-10691 Stockholm
- Norling-Christensen, Ole, ledende redaktør, mag.scient., f. 1940, Den Danske Ordbog, Københavns Universitet, Njalsgade 80, DK-2300 København S
- Poulsen, Sonja, lektor, f. 1948, Institut for Sprog og Kommunikation, Odense Universitet, Campusvej 55, DK-5230 Odense M
- Ore, Christian-Emil, førsteamanuensis, f. 1957, Universitetet i Oslo, Dokumentasjonsprosjektet, Boks 1102 Blindern, N-0317 Oslo

- Ralph, Bo, professor, f. 1945, Göteborgs universitet, Institutionen för svenska språket, S-41298 Göteborg
- Rode, Eva, ordbogsredaktør, f. 1945, Ordbog over det norrøne prosasprog, Njalsgade 76, DK-2300 København S
- Sanders, Christoffer, ordboksredaktør, f. 1948, Ordbog over det norrøne prosasprog, Njalsgade 76, DK-2300 København S
- Schoonheim, Tanneke, redaktør, f. 1965, Vroegmiddelnederlands Woordenboek, Instituut voor Nederlandse Lexicologie, Postbus 9515, NL-2300 RA Leiden
- Schutz, Rik, forlegger, f. 1953, Van Dale Lexicografie vb, Postbus 19232, NL-3501 DE Utrecht
- Seppinen, Hannele, frilancer/forskarstuderande, f. 1953, Rödbergsgatan 16 C 50, SF-00150 Helsingfors
- Simensen, Erik, førsteamanuensis, f. 1938, Universitetet i Oslo, Avdeling for leksikografi, Boks 1001 Blindern, N-0315 Oslo
- Svensén, Bo, handsekreterare, f. 1941, Svenska Akademien, Box 2118, S-103 13 Stockholm
- Svensson, Lars, ordbogsredaktør, docent, f. 1938, Svenska Akademiens ordboksredaktion, Box 3, S-221 00 Lund
- Sørensen, Viggo, lektor, mag.art., f. 1941, Aarhus Universitet, Institutt for Jysk Sprog- og Kulturforskning, Niels Juelsgade 84, DK-8200 Århus N
- Tarp, Sven, kandidatstipendiat, f. 1952, Handelshøjskolen i Århus, Fuglesangs Allè 4, DK-8210 Aarhus V
- Toporowska-Gronostaj, Maria, fil.lic., f. 1945, Göteborgs universitet, Institutionen för svenska språket, S-41298 Göteborg
- Viberg, Åke, docent, f. 1945, Stockholms universitet, Centrum för tvåspråkighetsforskning, S-10691 Stockholm
- Widell, Peter, ordbogsredaktør, cand.mag., f. 1945, Politikens Forlag A/S, Vestergade 26, DK-1456 København K





