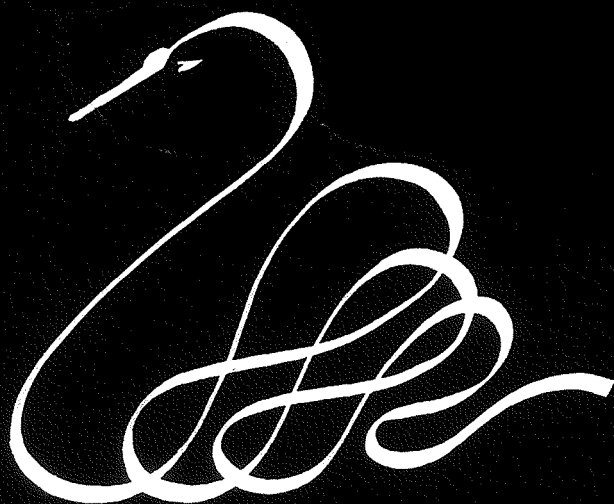


NORDISKE STUDIER I LEKSIKOGRAFI II

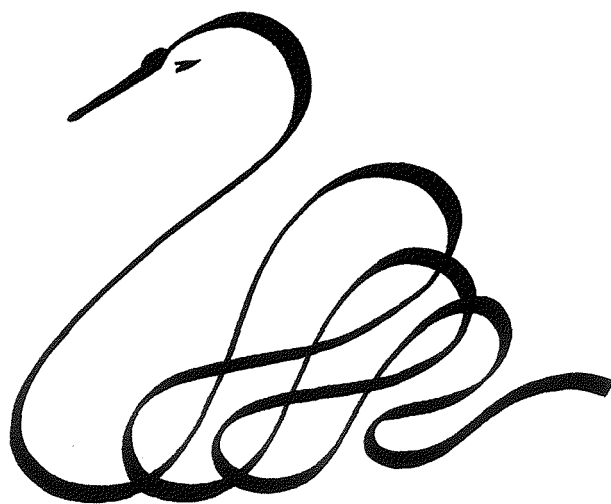


Skrifter udgivet af Nordisk Forening for Leksikografi
Skrift nr. 2

NORDISKE STUDIER I LEKSIKOGRAFI II

Rapport fra Konference om Leksikografi i Norden
11. - 14. maj 1993

Redigeret af
Inna Garde og Pia Jarvad



Skrifter udgivet af Nordisk Forening for Leksikografi
Skrift nr. 2

©LEDA (Foreningen af Leksikografer i Danmark)

Sats: Anna Garde & AKA-PRINT A/S, Århus
Tryk: AKA-PRINT A/S, Århus

ISBN 87-984909-0-7

Øvrige titler i serien:
Nordiske Studier i Leksikografi, Oslo 1993

Bogen kan rekvireres gennem:
Nordisk Språksekretariat
Postboks 8107 Dep
N-0032 Oslo

Indledning

Den anden Konference om Leksikografi i Norden blev afholdt i København i dagene 11.-14. maj 1993. Denne rapport indeholder de fleste af de foredrag der blev afholdt på konferencen - kun én deltager har ikke ønsket sit bidrag publiceret, men til gengæld har to, som var forhindret i at holde deres foredrag, indsendt dem til rapporten, hvilket vi anser som et stort plus.

Konferencen blev arrangeret af LEDA (Foreningen af Leksikografer i Danmark) i samarbejde med Nordisk Språksekretariat. Konferencekomiteen bestod af hovedparten af LEDAs bestyrelse, nemlig:

Jens Axelsen, Gyldendal
Anna Braasch, Center for Sprogteknologi
Anna Garde, G.E.C. Gads Forlag
Ebba Hjorth, Den Danske Ordbog
Pia Jarvad, Dansk Sprognævn

Redaktionskomiteen for rapporten har bestået af Anna Garde og Pia Jarvad.

Arrangørerne af konferencen havde fra første færd valgt ikke at opsætte nogen faste rammer for hvilke emner der kunne bringes på bane, omend vi i første rundskrivelse påpegede det ønskværdige i at nogle af de emner der savnedes ved 1. konference, blev taget i betragtning denne gang. Det gjaldt fx ordbøger som kommercielle produkter og leksikografiens historie i Norden. Det sidste emne blev da også taget op af nogle af deltagerne.

Foredragene kom til at spænde meget vidt og var så forskelligartede at redaktionen valgte at undgå en emneinddeling af rapporten og i stedet tyede til den for leksikografer naturlige alfabetiske orden - i denne rapport ordnet efter forfatterens efternavn. Bag i bogen er der et stikordsregister som er kommet i stand ved at de enkelte bidragydere har markeret de relevante stikord i deres korrektur, og i de få tilfælde hvor forfatteren ikke havde foretaget markeringer har redaktionen gjort det. I et særskilt register kan man finde de ordbøger der er blevet omtalt i foredragene.

Konferencen modtog støtte fra Statens Humanistiske Forskningsråd, Nordisk Kulturfond og Clara Lachmanns Fond, og Nordisk Språksekretariat har ydet støtte til udgivelsen af denne konferencerapport. Munksgaards Forlag bidrog med vinen til åbningshøjtideligheden, og G.E.C. Gads Forlag var vært ved afskedsreceptionen.

København, april 1994

Anna Garde og Pia Jarvad

Indhold

- Applicering af alordbogsprincippet – Sandro Nielsen 204
- Betydningsbeskrivelser – Ebba Hjorth og Knud Troels Thomsen 108
- Blant fire millioner sedler – Christian-Emil Ore 243
- Datamatstøttet vending af ordbøger – Anders Drejer Nygaard 237
- Den Danske Ordbogs tekstkorpus – Kjeld Kristensen 138
- EGs ord och uttryck – Anne-Lena Bucher 38
- Ekvivalensproblematikk – Ruth Vatvedt Fjeld 78
- Genbrug af ordbogsdata – Dorthe Duncker og Ole Norling-Christensen 61
- Halvfasta fraser – Erika Lyly 173
- Idiom och variation – Ulla Clausén 47
- Kontrastiv metaforik – Kerstin Norén 227
- Leksikografi og excentricitet – Jens Axelsen 18
- Lemmasektion til NLO – Hans Kristian Mikkelsen 193
- Lexikografins terminologi – Bo Svensén 263
- Mot en mångdimensionell tvåspråkig ordbok – Maria Toporowska Gronostaj 96
- Nogle generelle problemer – Karl Hyldgaard-Jensen 123
- Når maskinen tager én på ordet – Anna Braasch 29
- ODS-Supplementet – Sv. Eegholm-Pedersen 69
- Om datering af ordene i dansk – Anne Duekilde 53
- Ordbog og encyklopædi – Karen Margrethe Pedersen 248
- Ordbogsredaktion med COMPULEXIS – Anders Geertsen 87
- Ordboksredaktør i dag – Inger Hesslin Rider 257
- Ordsamlinger på dialekt – Olaf Almenningen 9
- Parsing af ikke-strukturerede ordbogsdata – Claus Bo Jørgensen 130
- Polylemi hos svenska verb – Åke Viberg 279
- Sammansättningsmorfologi – Sven-Göran Malmgren 179
- Skandinavisk ordbok – Birgitta Lindgren 159
- Strejflys over islandsk leksikografis historie – Gudrun Kvaran 152
- Svensk ordbok 1948 – Carl-Erik Lundbladh 166
- Vart tog Svensson vägen? – Kristina Nikula 218
- Översättarhjälp i en ordbok – Nina Martola 185
- Oversigt over forfatterne 291
- Ordbogsregister 293
- Stikordsregister 299

Ordsamlinger på dialekt

- problem og utfordringar

Olaf Almenningen

I arbeidet med *Norsk Ordbok* - ordboksverket over nynorsk skriftspråk og norske dialektar - har vi heile tida hatt stor nytte av lokale ordsamlinger og innsamla dialektmateriale frå større område eller landsdelar. Også vitskaplege utgreiingar om lydverk, formverk eller syntaks frå eit einskildmålføre har verdi for *Norsk Ordbok*, for ikkje å snakka om samlingar av ord og omgrep frå ymse fag og yrke, dyre- og planteliv.

Arkivet vårt er bygt opp av to hovuddelar: ein skriftleg-litterær nynorskdel og ein målføredel. Til målføredelen høyrer òg eit nett av vel 1000 lokale medarbeidarar som frivillig har sendt inn einskildopplysningar om ord og seiemåtar på setlar og eigne lister opp gjennom åra. Folkloristisk litteratur må òg nemnast i denne samanhengen, for her er mykje dialektprega skriftspråk, og somtid ofte eigne ordlister med i tilknytning til tekstane. Eigne samlingar av fyndord og ordtak må sjølv sagt heller ikkje gløymast.

Om lag 450 kjelder i *Norsk Ordbok* er reine skriftlege målføresamlingar eller -utgreiingar om ord og seiemåtar. Kring 300 (3/4) av desse att er i røynda dialektmonografiar. Dei 150 lokale ordsamlingane som då står att, er oftast ordrike kjelder som gjev stort tilfang til arkiva, reint kvantitativt sett, når dei blir ekserperte. Samla utgjer då denne delen av kjeldematerialet til *Norsk Ordbok* om lag 15% av alt grunnlagstilfanget i tittelstal. Vi viser elles til kjedelista i *Norsk Ordbok* for dei som er spesielt interesserte i dette.

Ordsamlingane er nokså ulike seg imellom og spreidde utover heile Noregs land. Dei fleste kjem nok helst frå bygdelag i Sør-Noreg. Somme er innsamla av ein einskildperson åleine, nokre av fleire i lag, andre att av skuleklassar, kulturorganisasjonar o.l. Ein del er av eldre dato, men svært mange av dei er faktisk nokså nye og vitnar både om aukande språkleg medvit og aktiv interesse for lokalkulturen i vår tid. Arbeidet med å samla inn ord og uttrykk har faktisk skote fart dei siste 10-15 åra.

Desse ordsamlingane kan variera mykje i storleik og omfang, frå eit par hundre oppslagsord og heilt opptil 20 000. Ragnhild Paulsens store ordsamling frå Nøtterøy i Vestfold (nr 4) er den største vi har til no. Dei fleste har likevel teke med ein stad mellom 1 000 og 3 000 ord i alt. Som regel set føremålet med ordsamlinga også opp rammene og omfanget.

Mesteparten av ordsamlingane inneheld allmennord i dialekten, men somme kan òg omtala særskilde ordgrupper, som t.d. kalleord eller skjellsord, fugle- eller blomenamn, ord for fiske og sjøfart, lokale båtnemningar, jakt- og fangstuttrykk eller byggjeterminologi. Det er samlingane av *allmennord* eg fyrst og fremst vil ta opp i denne artikkelen,

men dei fleste metodeproblema eg kjem til å nemna, gjeld i like høg grad andre typar lokalmerkt språktilfang.

Problema står eg sjølv midt oppe i, då eg i to år no har arbeidd saman med ein heimsmann frå Sauda i Ryfylke med å få skipa til ei ordsamling frå dialekten der. Eg har valt ut 12 ulike samlingar av ymsande slag i denne framstillinga, spreidde så nokolunde likt utover landet, men likevel mest frå Sør-Vestlandet, der eg sjølv er best kjend.

Dei viktigaste leksikografiske metodespørsmåla ser ut til å vera: 1) avgrensing av ordutval, 2) val av oppslagsform, 3) uttale og grammatisk merking og 4) tyding(ar) og eksemplmateriale.

1) Kva ord skal takast med i ei lokal ordsamling?

Når ein skal setja saman ei lokal ordsamling, kjem ein fort opp i avgrensingsproblem: Kva slags ord fortener ein plass, og kor langt skal ein gå i å ta med allmenne, vidspreidde ord som ikkje plent høyrer heime berre i eitt bygdelag eller einast i ein by? Dei fleste orda våre gjer jo ikkje det. Dessutan er det slett ikkje så sikkert at eit ord som kjennest svært lokalt ut, eigentleg er så lokalt når det kjem til stykket. Professor Egil Pettersen drøftar dette utvalsproblemet i forordet sitt til *Bergensordboken* (nr 2), der han mellom anna seier dette:

"Hva skal vi så regne som bergenske ord? (...). Prinsipielt skulle det være alle ord som blir brukt i bergensk, også alle de felles norske ordene, som nok er de helt dominerende. Men det er ikke hensikten med en bok som denne å ta slike ord med. Skal vi så avgrense oss til ord som Bergen er alene om? Det ville gi et magert utbytte. For en stor prosent av de ordene vi kjenner som typisk bergenske, er kjent og brukt også andre steder, først og fremst på Vestlandet og i Nord-Norge. Men ord som ikke er allment brukt i norsk, som altså ikke hører med til bokmålets ordforråd, må få plass i en bergensordbok". (s. 9)

Tilhøvet til det nasjonale skriftmålet må såleis ikkje spela noka avgrensande rolle for utvalet. Forfattaren skal fyrst og fremst vera trugen mot det lokale språket og gje det rett og rom.

Hjå *Johan Hveding* i hans *Håløygsk ordsamling* (nr 12) frå Nordland og Troms har vi dertil eit par utvalskriterium som viser ei noko anna side ved dette problemet. Han nemner desse grunnane til å ta med eit ord i samlinga si:

- når ordet ikkje er kome med i ordbøker før,
- når eit ord vert brukt i ei anna tyding enn det som tidlegare har vore oppført,
- når eit ord berre er heimfest til Sør-Noreg i ordbøkene
- når ordet er allment nordpå, medan det vert rekna for sjeldan i sørluten av Noreg. (frå føreordet)

Ei sterk og medviten framheving av det nordnorske ordtilfanget er det viktigaste føremålet med denne ordsamlinga. Regional sjølvhevdning er eit viktig program for forfatta-ren. Han vil fylla ut det norske riksspråket med regionalt måltilfang. Nordnorske dia-

lektord skal lyftast opp på nasjonalt nivå og stå i jambreidd med sørnorske ord.

Det er som regel det meir sjeldsynte og særmerkte ordtilfanget dei fleste ordsammarane vil ta med og visa fram. I ordsamlingane frå dansketida (jf *Thomas Bloch: Fyresdal 1698*, nr. 1) var dette mest den einaste grunngevinga. Innsamlingane frå denne tidbolken har helst med berre einskildord, mindre med faste ordelag eller lokale idiom. Men når allmenne ord er tekne med, er det som oftast i tilknytning til ei særskild vending frå målføret i bygdelaget.

Ei kjensle av og redsle for orddød, språkleg gløymse - og gjerne forfall - pregar somme av fororda i dialektordsamlingane. I tillegg kjem viljen til å visa språkleg rikdom og fantasi, mot meir moderne einsretting og forflating. I ordtilfanget skjer det gjerne snøge endringar, der nasjonale ord trengjer ut meir lokale nemningar og uttrykk. Titlane kan stundom bera bod om dette, som t.d. Daniel Dvergsdals *Gamle ord frå Jølster* (nr.9). Han grunnjev sitt utval med at:

“Siktemålet vart å skrive ned ord som er gløymde av yngre jølstringar eller er i ferd med å verte utbytte, i alle fall i skrift. Ein del av desse orda er knytte til gammalt arbeidsliv og gamle skikkar. Andre er godt brukande i vårt moderne samfunn for den som vil halde på målføret sitt og elles legg vekt på presis og variert ordbruk”. (frå føreordet)

Og *Olav Hetland* frå Hjelmeland i Ryfylke (nr. 5) illustrerer sitt ordsamlingsprosjekt med dette råkande biletet:

"Ordtilfanget i dialekten kan samanliknast med eit vatn. Frå bekkjer og elvar får vatnet stadig nye tilførsler. Samstundes er det ein jamn straum som renn ut. Det er denne utgåande straumen me tek for oss i denne boka" (s. 7).

Det "unormale" og "avstikkande", attåt det "utdøyande" og "tapande" har altså ein framskoten plass, særleg i dei minste samlingane. I større samlingar har også allmennorda eit breitt rom, såleis i *Oppdalsmålet* av *Ingeborg Donali* (nr. 3) og *Ordbok over nøttlands-målet* av *Ragnhild Paulsen* (nr 4).

Ønsket om å visa og ta vare på eit språkleg mangfald lokalt stikk altså djupt hos mange av ordsammarane våre. Eit anna føremål er somtid å gje dei yngre generasjonane forklaringsringar på somme av dei mest avstikkande orda og uttrykka i heimemålføret. Dei veit gjerne ikkje så heilt nøyaktig kva desse orda tyder lenger, og treng å bli minte om sine språklege røter.

Vi bur jo i eit langstrekt og griskrendt land, som alle veit, med spreidd busetnad og ei nokså sterk regional og lokal identitetskjennele. Både dette faktum og den positive språkinteressa som følgjer av den norske målstriden, gjev rimeleg gode vokstervilkår for all lokal ordsammarverksemd. Det reint deskriptive perspektivet åleine vantar hos dei fleste ordsammarane, om det då i det heile er mogeleg å vera heilt verdifri og berre registrerande i slike sosiale spørsmål som har med språk å gjera. Ordutvalet er gjerne nøye gjennomtenkt og vurdert. Dei lokale ordsammarane har slett ikkje noka interesse av å tona ned eller løyna føremålet med sammararbeidet sitt. Dei er normative og vil svært gjerne vera det.

2) Oppslagsform

Dei største problema med lokale og regionale ordsamlingar gjeld sjølve *oppslagsforma*. Her finn vi då òg dei største variasjonane. To hovudtendensar synest likevel å peika seg ut:

- normert språk (oftast nynorsk) i oppslagsforma
- dialekt i oppslagsordet, gjerne så nær den lokale uttalen som råd.

Sidan det er det reint lokale som har størst interesse i slike leksikografiske prosjekt, er det viktig å gje att lokalpreget på eit eller anna vis. Orda må presenterast på ein like heimleg måte som dei vert brukte. Og vi ser då òg at dei som har valt nynorsknormerte oppslagsformer, gjerne går eit godt stykke i dialektlei og nyttar ut nettverket av sideformer. Stundom bryt dei beintfram med rettskrivingsreglane, særleg der ordet skil seg avgjerrande ut frå normalen (t.d. *juahus* i *Sandvik: Gamle ord frå Suldal*, nr 7).

Ein skulle i fyrste omgang tru at reine målføreoppslag, attgjevne på ekte dialekt måtte vera rettaste vegen å gå når det heimlege skal stå i høgsetet. Men her støyter ein på vanskar som har med språkvanar og ordbokskonvensjonar å gjera. Folk flest tykkjer det er vanskeleg å lesa dialekt, også sin eigen. Og for folk utanfrå, blir desse problema berre endå større. Dessutan kjem vi opp i problem når vi skal gje att dialektuttale direkte i skrift. Skal vokallengd markerast, og korleis skal vi gje att lange konsonantar og vokal-bortfall i innlyd t.d.? Og kva med tonelag, sløyfing av konsonantar i innlyd og ending, tjukk og tynn *l*, diakritiske teikn o.l.? Og korleis skal skriftbiletet sjå ut: Skal vi velja vanlege bokstavar eller lydskrift? Og kva lydskriftsystem skal vi i så fall bruka, om vi vel *den* utvegen?

Dei som har valt denne løysinga (jf *Hetland* nr 5, *ƒenshus* nr 6 og *Time* nr 8), gjev dermed lesarane sine ein del problem med identifikasjonen av oppslagsorda. Vi må gå ut frå at mange av dei ikkje har nokon avansert språkvitskapleg bakgrunn og dermed er lite kjende med dei ulike lydskriftsystema. Kva slags *ord* er *rå:d'n* hjå *Sandvik* eller *je:re* hjå *Time* t.d.?

Skrivemåten av ymse konsonantar i framlyd valdar største problemet. Det er lettare å både lesa og skjøna kva for eit ord vi talar om når vi ser stavemåten *gisen* i normert form framfor *ji:s'n* i dialektskrift. Andre tider finn ikkje folk eit ord att, fordi det ikkje står på den alfabetiske plassen der ein hadde venta å finna det (jf *jot* for *ljot*). Her kan altså eit ord rett og slett bli "borte" for lesaren. Dette gjeld til dømes for ord med konsonantsamband i framlyd (særleg *j*-variantane), og noko avstand mellom uttale og skriftbiletet.

Folk leitar gjerne ut frå norma, mellom anna fordi dei sit med eit "tenkt" skriftbiletet inni seg på førehand. Om ein vel den lydmerkte og dialektneare oppføringsmåten, bør ein iallfall vera påpasseleg med krysstilvisingar (jf *Sandvik*, nr 7). Dette saknar eg blant mange av dei ordsammlarane som har valt denne metoden. Flestalle i denne gruppa har likevel brukt vanlege bokstavar framfor fonetiske spesialteikn og avvikande teikn for ymse vokal- og konsonantvariantar. Dette lettar både lesing og forståing for vanlege folk.

Faren med normering til normalmål er fyrst og fremst at mykje av den lokale svipen vert borte. Slik tykkjer eg det er langt på veg i ordsamlinga *Kvinnheringskringlo* (nr.10),

der både oppslagsord og døme er normerte til nokså fast nynorsk normalmål (1917-rettskriving). Men sjølv sagt har det òg sin verdi å vita kva det lokale ordet heiter i normalmålet.

Somtid kan vi koma opp i språkstrid også, særleg i landsdelar der nynorsk er lite brukt. Folk er for dialekt, men uvande med nynorsk. I slike tilfelle har ordbokssamlaren då gjerne valt den reine dialektløysinga (jf *Gamle ord frå Rødnes*, nr 11) eller ei blanding. Andre gonger kan det vera usemje om sjølv norma, t.d. ein moderne variant av nynorsk (1959-normalen) mot ein eldre variant (1917-rettskrivinga).

Ein annan utveg er å la oppslagsorda stå på tilnærma dialekt, medan ordforklaringane har normalform (jf *Jenshus: Fronsmålet*, nr 6). Uttaleproblemet er då løyst saman med oppslagsforma, og ein har fått *eitt* samla oppslag i staden for *to* oppføringar. Men problemet er der likevel dersom ein går for langt i fonetisk lei med skrivemåten, slik at identifikasjonen vert borte. Somme har her valt smidige mellomveggar, t.d. i *Bergensordboken* (nr. 2), der vi kan lesa både bokmål, nynorsk og bergensdialekt ut av dei normerte oppslagsorda (*fryntelig, handlaus og kisøge*). Dette er nok gjort på denne måten, rett og slett fordi orda er slik og heiter så i bergensmålet. Rein og konsekvent normering til standardbokmål av dialektord finst det lite av i våre kjelder, heller ikkje bymåla vert handsama såleis.

De skjønar sikkert at eg sjølv hallar mest til løysinga med romsleg normerte oppslagsord og lokal svip i alt dømematerialet. I Sauda-prosjektet har vi landa på den framgangsmåten, med mange småjusteringar attåt. Det er om lag slik *Sigurd Sandvik* òg har gjort det i *Gamle ord frå Suldal* (nr 7). Men normeringa må gjerast varsamt og med skjøn, og ut frå dialekten sjølv, så langt det går. Normeringa bør syna samanheng med dialekten og ikkje visa bort frå han. Dessutan er det lurt med mange interne krysstilvisingar på alfabetisk plass, særleg for ord som sprikjer mykje i uttale og skriftbilete, og elles for ord som ikkje er særleg godt kjende i skrift, men berre i talemålet. Dette har vi nytta oss mykje av i *Norsk Ordbok*.

Ein kan nemleg ikkje venta at lesarane skal slå opp på normert plass alltid (jf problemet med ordet *kjuagutt* vs *tjuagutt* i bergensmålet), særleg viss dei kjenner ordet som reint talemål og går ut frå det når dei skal leita det opp. Vi kan ofte støyta på ord som jamvel er berre å finna i talemålsform og ikkje står i dei største norske ordbøkene eingong (jf det vanlege nordnorske dialektordet, *jåssa* (= vasa, tøysa)). Skrivemåte og plassering av slike ord er då ikkje så heilt lett alltid.

Rein, dialektfri normering er etter mi meining derimot ikkje ønskjeleg i slike ordboksprosjekt, for då er noko av føremålet med ei slik ordsamling borte.

3) Uttale og grammatisk merking.

Dialektane og dei lokale språktradisjonane må sjølv sagt koma fram på ein sentral måte i ordsamlingar som vi her snakkar om. Spørsmålet blir i røynda *kvar* dette skal koma til uttrykk og *kvar* det gjer seg best, m.a. ut frå lesarane og mottakarane sine behov.

Somme har, som vi har sett, løyst dette ved å bruka dialekt i oppslagsformene (m.a.

Jenshus, nr 6, *Time*, nr 8, *Hetland*, nr 5 og *Dvergsdal*, nr 9). Andre har helst valt tilnærma normalmål i oppslaga, og late dialekten få meir eller mindre fritt spelarom i *brukseksempla* (t.d. *Donali*, nr 3, *Paulsen*, nr 4 og *Sigurd Sandvick*, nr 7). Ein tredje utveg kan gjerne vera å ha med eiga uttalemarkering, bak det normerte oppslagsordet. Det er noko slikt vi har i *Norsk Ordbok* når vi der gjev opp varierende målføreuttale i hakeparentesen bak oppslagsforma. Svært få lokale målføresamlingar har dette, dei fleste nøyer seg med å visa uttalen gjennom bruksdøma sine, om dei då ikkje har han i oppslagsforma. Dette kjem truleg av at dei kjenner uttalen som ein del av sjølve identiteten åt ordet, derfor skal det heller ikkje vera noka meir eller mindre tilfeldig tilleggsopplysning. I ordsamlinga frå Hjelmeland (nr 5) er til og med tonelaga markerte atmed oppslagsordet, der forfattaren har funne det nødvendig.

Når det gjeld ordsamlinga frå Sauda, har vi òg landa på denne siste løysinga. Vi kan då ta ut uttalen, om han likevel kjem tydeleg med i bruksdøma. Dette sparar plass og gjev likevel ordet att i sitt rette element. Norsk Ordboks heimelsmann har då fylt ut med bruksdøme for alle ord der dette mangla i det fyrste materialet vi fekk frå hans hand.

Grammatisk merking har dei fleste teke med der vi etter konvensjonen plar ha det, nemlig like bak oppslagsordet. Mange ordsamlingar har likevel lagt lite vekt på å ta dette med. I eldre ordsamlingar kan det vera så ymse med både grammatikken og oppslagsorda, såleis i *Fyresdal 1698* (nr 1), der verb står i presensform på oppslagsplass (Døme: *Baagnar*, *Flecto*, *Bøyer*), substantiv gjerne i bunden form (Døme: *Barkien*, *Guttur*, *Strube*) osv. Somme føyer gjerne på bøyning/morfologi attåt, dersom plassen tillet det (jf *Jenshus*, nr 6). Men elles er erfaringa gjerne den at uskolerte, men språkvakne folk stundom har problem med dette punktet, og det er gjerne her forskingsmiljøa kan vera til hjelp. Fleire av ordsamlarane våre har søkt støtte til denne sida av ordsamlingsarbeidet, og ordbøkene har vorte til i vekselverknad mellom heimelsfolk og lokalmiljø på den eine sida og forskingsmiljøa på den andre.

I tilfellet Sauda har eg såleis ordna den grammatiske delen ut frå døma åt heimelsmannen og spurt meg fram der uvissa rådde. Denne dialogen har vore sær s fruktbar, og har mellom anna ført til at nye nyansar i tilfanget har kome for dagen. Det har sjølv sagt interesse, særleg for målvitskapen, men òg for lesarar lokalt, å få greie på korleis eit verb blir bøyggt eller kva genus eit substantiv har i eit målføre. I normeringa av nynorsk er dette òg ei viktig rettesnor.

4) Tydingar og eksempelmateriale

Attåt uttale og lokal utsjånad elles, er det i hovudsak *tydinga* av orda som folk spør etter og er mest interesserte i. Dei vil ha sjeldsynte eller lokale ord *forklarte*, gjerne med meir vanlege og kvardagslege synonym, eller eigne utgreiingar dersom det er snakk om reiskapar frå gamalt arbeidsliv t.d. Det er jamnast òg denne delen av dialektordboka som får breiast omtale og mest utførleg plass (jf *Hålløygsk ordsamling*, nr 12 og *Bergensordboken*, nr 2). Tydingsdelen er *alltid* med, også der andre sider ved ordsamlinga manglar,

t.d. uttale eller grammatisk merking. Det seier seg sjølv at denne sida ved det lokale språkmaterialet er sær verdfull for den leksikografiske vitskapen.

Det er ein tendens til at definisjonane og synonyma jamnast er å finna i normalmåls form i denne delen. I dei gamle ordsamlingane frå dansketida brukte dei jamvel latin som forklaring på somme norske ord, i tillegg til dei danske definisjonane (*Bloch: Fyresdal 1698*, nr 1). Slike løysingar med metaspråk finn vi lite av i moderne ordsamlingar.

Eit ord er nake utan språkleg samanheng eller bruksdøme. Men det er likevel ikkje støtt at forklarings- og tydingsdelen berre står fram i form av synonym, utførlege definisjonar eller direkte omsetjingar åleine. Både eksempla og ordforklaringane *til saman* utgjer i grunnen definisjonen av eit ord, og i somme ordsamlingar er eksempelmaterialet så rikt, at den direkte forklaringa eller definisjonen mange gonger ikkje er med. Dette er særleg tilfellet i større samlingar som *Paulsen* (nr 4) og *Donali* (nr 3). Dette gjeld òg mest for kvardagslege ord, der meininga er godt kjend elles og eintydige definisjonar vanskelege å gje.

I kortare framstillingar er det likevel dei direkte tydingane som dominerer. Mange gonger kan slike "definisjonar" vera i knappaste laget, gjerne berre eitt einaste synonymt ord, som igjen kan vera fleirtydig i seg sjølv. Det er heller ikkje alltid slik at dei forskjellige tydingane er merkte av med ulike tydingsnummer, særleg gjeld dette i eldre ordsamlingar, der framstillinga gjerne er mangelfull, sett ut frå ein grammatisk og leksikografisk synsstad. Stundom kan jamvel homonym skrivast under eitt eller blandast i hop. Somme forfattarar skil heller ikkje alltid godt nok mellom ordklassane, og den språklege symmetrien mellom oppslagsord og forklaringsdel manglar somme stader.

I somme tilfelle er tydingsdelen grundig og svært fyldig, med tilvisingar til andre nabospråk og etymologi, såleis i *Bergensordboken*, der forfattaren følgjer eit ord i bruk bakover både i målføre og skjønnlitteratur. Våre ordsamlarar i Norsk Ordbok bruker ofte tilvisingar til *Hans Ross'* og *Ivar Aasens* ordbøker, og noterer om dei finn likskapar eller avvik i høve til dei opplysningane som står hjå dei.

Eg sa tidlegare at mange av ordsamlarane - og lesarane - er opptekne av det spesielle og avvikande ved tilfanget sitt. Det er då sjølvstekt ein fare for at ein del heimelsfolk kan hengja seg opp i ei særtyding som er *for* spesiell eller individuell. Slikt kan vi sjølvstekt aldri gardera oss imot, for det spesielle og fargerike er alltid meir spennande enn det allmenne. I slike tilfelle er det viktig at både særtydingar og meir vanlege bruksmåtar av eit ord kjem med, så biletet av bruksområdet for eit ord vert så breitt og rett som råd. Ein leksikograf må støtt vera på vakt mot det altfor eksotiske i slike samlingar.

Samanlikningar med dialekt- og ordsamlingar frå grannebygder eller nabokommunar kan vera *ei* korrigerande kjelde og rettesnor, likeins jamføringar med andre heimelsfolk på staden, om vi er så heldige at dei finst. Det vil sjølvstekt ikkje seia at vi ikkje stolar på våre heimelsfolk, men heller at vi ønskjer å få vita *meir* om alt det dei veit, men som dei ikkje trur det er noko særleg å bry seg med. Det kan mange gonger sjå så altfor sjølvstekt ut for dei. Heilt fullvisse på at vi har fått fram det vesentlege kan vi, prinsipielt sett, likevel aldri bli, så her må vera eit breitt rom for sakleg skjøn og audmjuk tilnærming.

Konklusjon

Å skipa til ei dialektordbok er fyrst og fremst eit leksikografisk arbeid. Eit nært samarbeid med dialektologane er òg nokså sjølvsgt. Og i meir avgrensa ordsamlingar innanfor snevrare fagfelt, som t.d. båtbygging eller plantenamn, vinn vi sjølvsgt på å ha med vurderingar frå kunnige fagfolk frå desse fagområda.

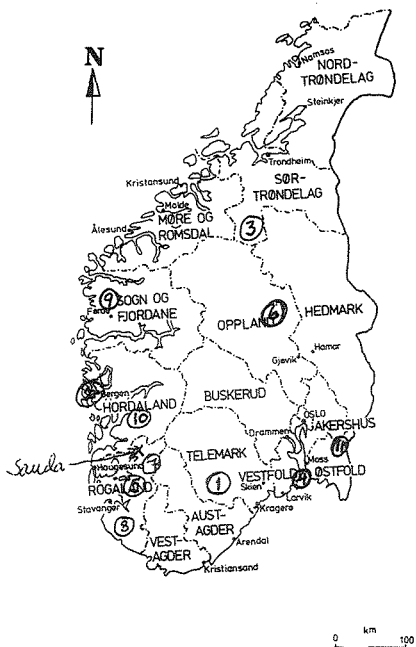
Men utfordringane er mange, og handverket er nokså ulikt utført mange stader. Det viktigaste må likevel støtt vera at dette arbeidet verkeleg vert *gjort*, og at folk samlar inn og legg fram det dei sit inne med av språklege kunnskapar og røynsler. Det er sjølvsgt ikkje meininga at alle *skal* skriva slike ordsamlingar heilt på same måten og etter same leisten, men det har noko å seia at hovudlinjene i ordboksredigeringa er følgde så godt som råd. Her spelar pedagogiske omsyn sjølvsgt ei sentral rolle. Likeins må arbeidet gjerast så heile språkmaterialet kjem til sin rett.

Elles er føremålet med slike lokale og regionale ordsamlingar sjølvsgt mest avgjerande for det endelege resultatet: Kven er lesarane og mottakarane? For oss leksikografar er det likevel viktig å samordna det arbeidet som vert gjort noko meir, slik at ettertida - både vitskapen og lesarane - kan få endå meir utbytte av desse samlingane. Det er jamvel ei oppgåve i seg sjølv å skaffa seg oversikt over dei lokale prosjekt som alt finst eller vel er komne i gang. Andre gonger kan vi fungera best som fødselshjelparar og gje tildriv til slikt arbeid. Som filologisk ekspertise har vi - saman med dialektologane - ein viktig funksjon som fagkonsulentar, slik at lekfolk og andre kan venda seg til oss med spørsmål om korleis dei skal ordna stoffet sitt. For det står fast at dette arbeidet er ei viktig kartlegging av heile språket vårt, så vi kan få fram den breidd, djupn og rikdom som der finst.

Litteratur

- 1 Thomas Bloch: *Glossemata Tellemarchia*. Fyresdal 1698. Oslo 1956.
- 2 Egil Pettersen: *Bergensordboken*. Bergen 1991.
- 3 Ingeborg Donali: *Oppdaling. Ord og uttrykk*. Trondheim 1988.
- 4 Ragnhild Paulsen: *Ordbok over nøttlandsmålet omkring 1900*. Tønsberg 1981.
- 5 Olav Hetland: *Eg seie meg så. Ord og uttrykk frå Hjelmeland*. Hjelmeland Mållag 1989.
- 6 Gunnar Jenshus: *Fronsmålet*. Vinstra 1986.
- 7 Sigurd Sandvik. *Gamle ord frå Suldal*. Suldal 1986 og 1991.
- 8 Tor Time: *Haurt saure*. Time Mållag 1986.
- 9 Daniel Dvergsdal: *Gamle ord frå Jølster*. Jølster 1990.
- 10 Anders Haugland: *Kvinnheringskringlo*. Kvinnherad 1978.
- 11 Ragnar Elvén Myhrvold og Sjur Wergeland-Krog: *Ord og uttrykk frå Rødenes*. Fredrikstad 1992.
- 12 Johan Hveding: *Håløygsk ordsamling II*. Bodø 1968.
- 13 *Norsk Ordbok I-II (a-flusken)*. Oslo 1950-1978.

Ordsamlinger på dialekt



Leksikografi og excentricitet

Jens Axelsen

På et af de "morsomme" skilte man kan købe til ophængning på kontorer står der: "Man behøver ikke være tosset for at arbejde her, men det hjælper." Jeg har sommetider spekuleret på om denne sentens også kan bruges om leksikografi: "Man behøver ikke være tosset for at dyrke leksikografi, men det hjælper", for når man ser tilbage i historien finder man adskillige eksempler på excentriske leksikografer: vor store forgænger **Dr Johnson** var bestemt ikke helt almindelig, og også i hans leksikografiske arbejde er der glimt af excentricitet. Velkendt er hans definition af *lexicographer*: "A writer of dictionaries, a harmless drudge", og af *oats*: "A grain which in England is generally given to horses, but in Scotland supports the people". Min yndlingsdefinition fra ham er dog følgende, af ordet *dogsbolt*: "Of this word I know not the meaning, unless it be..." og så kommer han med en lang og skrupforkert forklaring.

En anden stor forgænger, **Noah Webster**, havde den ide at han ville lave sproget om så det passede til hans opfattelse af hvordan det logisk burde være (og han havde også held til at få ændret nogle stavemåder i amerikansk engelsk). Han syntes nu at når *ea* udtaltes /i:/ i *leaf* og *sheaf*, burde *d e a f* også udtales /di:f/ selvom det faktisk hedder /def/ så han rettede hele sit liv folk der sagde /def/ i hans pãhør.

Og mit emne her er en tredje leksikograf, der ganske vist ikke har samme status i leksikografiens historie som de to nævnte koryfæer, men til gengæld nok langt overgår dem i excentricitet, nemlig islendingen **Thorleif Gudmundsson Repp**.

Hans interesse i en leksikografisk sammenhæng er den, at han i 1845 sammen med englænderen *James Stephen Ferrall* udgav en Dansk-Engelsk Ordbog der blev et vendepunkt i dansk-engelsk leksikografi. Med den forlod man d. 18. årh.'s tradition, som de hidtidige dansk-engelske ordbøger var præget af, og fik en mere moderne ordbog. Den var den første der havde et solidt fundament at bygge på, såvel for de danske ords vedkommende, nemlig *Molbechs* "Dansk Ordbog" fra 1833, som for oversættelsernes, da Ferrall var englænder, og Repp selv var særdeles velbevandret i engelsk. Bogen har stadig en efterkommer her i landet, nemlig Gyldendals røde dansk-engelske ordbog, der i lige linie, gennem forskellige udgivere, nedstammer fra den, og det var også det der var min tilgang til Repp, da bogen jo således i 1995 har 150 års jubilæum. Jeg anede ikke hvem Repp var da jeg så hans navn på titelbladet til ordbogen, og ingen jeg spurgte kendte ham. I *Dansk Biografisk Leksikons* sidste udgave er han ikke med (han var dog med i de to foregående udgaver). Han er faktisk nu næsten fuldstændig glemt her i landet. Men da jeg forfulgte hans spor i samtidige memoirer, i breve og i danske arkiver, kom der interessante ting for dagen, som det vil fremgå af det følgende. Forøvrigt er der

tegn der tyder på en fornyet interesse for ham: i Island blev der i 1989 rejst et mindesmærke for ham på kirkegården i Reykjavik, hvor han ligger begravet. Ligeledes i Reykjavik, hvor en stor del af hans efterladte papirer findes, endnu uordnet, på Landsbiblioteket, er der også for nylig (1991) udkommet en bog om ham skrevet af en engelsk forsker fra Leeds *Andrew Wawn*. Den omhandler hovedsagelig hans tid i Edinburgh, som de danske kilder næsten ikke omtaler.

Thorleifur Gudmundsson Repp var født på Island 1794, studerede 1814-26 i København, hvor han blev særdeles velbekendt i akademiske og litterære kredse, var bibliotekar ved *The Advocates' Library* i Edinburgh 1826-37, og vendte derefter tilbage til København, hvor han døde 1857.

Det er ikke Repps ordbog der er mit emne ved denne lejlighed, da det ikke er her excentriciteten viser sig, og heller ikke hans andre værker, men hans personlighed. Jeg vil koncentrere mig om fire episoder:

- 1 Hans forbindelse med Bakkehuset, hvor den kendte professor i litteraturhistorie *Knud Lyne Rahbek* sammen med sin livlige og begavede ægtefælle *Kamma* havde en slags litterær salon, hvor tidens førende kunstnere kom.
- 2 Hans magisterdisputats, der fik et meget særpræget forløb.
- 3 Hans stormfulde tid som bibliotekar i Edinburgh.
Og endelig
- 4 hans ejendommelige kontrovers med Danmarks verdensberømte fysiker og opdager af elektromagnetismen *H.C. Ørsted*, efter tilbagekomsten til København.

Disse ret underholdende, men også på en måde tragiske, aspekter af hans liv er selvfølgelig ikke hele sandheden om Repp. Han var en højt begavet mand med et exceptionelt nemme for sprog: Han kendte foruden latin og græsk, som han allerede udmærkede sig i i skolen, oldislandsk (naturligt nok), oldengelsk, gælisk, hebræisk, arabisk, finsk og ungarsk samt de vesteuropæiske sprog: engelsk, tysk, fransk, spansk, italiensk, portugisisk. I nogle af sine bøger henviser han også til sanskrit og eskimoisk. Hans lærdom spændte vidt, også inden for andre områder: Han læste de første to år medicin med særlig henblik på dens forbindelse med kemi og fysik (som han studerede under *Ørsted*), gik senere over til de humanistiske fag: filosofi, litteratur, historie, filologi, og blev stærkt optaget af den nye sprogvidenskab under indflydelse af sprogforskeren *Rasmus Rask*. Repp vandt guldmedalje for afhandlinger i filosofi og æstetik 1818 og 1823, han oversatte 1825 "Laxdøla Saga" til latin for den arnamagnæanske kommission, der forestod (og forestår) udgivelsen af de islandske håndskrifter som *Arni Magnussen* indsamlede på Island i begyndelsen af 1700-tallet, og han vandt europæisk berømmelse med sin bog: "Historical Treatise on Trial by Jury... in Scandinavia and in Iceland" udgivet i Edinburgh 1832. Han oversatte bøger og skrev afhandlinger og avisindlæg om et utal af emner inden for historie, runologi, klassisk filologi, teologi, medicin, landbrugsvidenskab, politik, som det fremgår af fortegnelsen over hans værker i *Erslews forfatterleksikon*. Desuden var han hele sit liv fuld af ideer og projekter, og hans virkelyst var utrolig lige fra hans stu-

dietid: han var medstifter af Studenterforeningen i København 1820 og med til at grundlægge læseselskabet "Athenæum" 1824. Han havde i Edinburgh planer om udgivelse af tidsskrifter, stiftelse af lærde selskaber, en forskningsrejse for at studere "the Tartarian and Caucasian tongues", og, som noget af det mere bizarre, oversatte han for sin egen fornøjelse Lukas og Johannes evangelierne til oldengelsk. Han holdt foredrag, underviste, anmeldte bøger. I sin anden periode i København var han for en tid redaktør af den da kendte avis "Dagen" og udgav senere sit eget tidsskrift "Tiden". Men omfanget af hans interesser var også en hæmsko for ham, og han havde tit vanskeligt ved at gøre ting færdige, hvad hans efterladte papirer vidner om som Wawn påviser. Finnur Jonsen skrev i artiklen om ham i Salmonsens leksikon (1926): "Han var en Mand med rige Evner, et skarpt Hoved, men spredte sig for meget."

Han færdedes blandt tidens kendteste skikkelser og havde en næsten H.C.Ander-sensk evne til at vinde betydende personer for sig: I Danmark bl. a. Knud Lyne og Kam-ma Rahbek, digteren og filosofen F.C.Sibbern, Rasmus Rask. På en rejse til England 1820-22 stiftede han bekendtskab med the Marquis of Londonderry, udenrigsministe-ren Lord Castlereagh, Lord Ellenborough, foruden en David Ker, der som svigersøn til marquis'en og svoger til Castlereagh skaffede ham disse forbindelser. Efter hjemkom-sten fra denne rejse blev han engelsklærer for prinsesse Caroline Amalie, gift med den senere kong Christian VIII, som ordbogen for øvrigt er dediceret til. Men han var også i besiddelse af en umådelig selvbevidsthed og et ustyrligt, ja Hekla'sk temperament (P.H.Boye siger i sin bog om Kamma Rahbek at "hans Fædreøs paa eengang haarde og ildsprudende Natur afspejlede sig i hans Væsen"). Dette temperament bragte ham ofte på u-Andersensk vis i heftig konflikt med omgivelserne, og det ganske uden persons an-seelse, som det vil fremgå af det følgende.

Gennem sit studium kom Repp som nævnt i forbindelse med *Knud Lyne Rahbek*, blev hans ven, underviste ham i islandsk og indførte ham i sagalitteraturen. Han leverede og-så oversættelser af islandske sagaer til Rahbeks tidsskrift "Dansk Minerva". "Fra Stude-restuen avancerede han til Fruens Cabinet" fortæller P.H.Boye. Folkemindeforskeren J.M.Thiele skriver i sine erindringer om gæsterne på Bakkehuset: "Islænderen, Dr. Repp var en sjældnere Fugl og ikke meget omgængelig paa Grund af sit noget hovmo-dige Væsen, der skaffede ham Navnet "Lord Repp", og under dette Navn - taaltes han." I et af sine breve til Kamma Rahbek (19 sept. 1821) spørger digteren Chr. Winther til "Lord min Sandten Repp", hvad det så skal betyde. Det kgl. Bibliotek har en del breve fra Repp til Kamma Rahbek, også citeret i Boyes bog, og det fremgår af disse at han har været en temmelig besværlig gæst, som hun flere gange måtte irettesætte for hans ube-herskede optræden over for andre gæster, heriblandt selveste *Oehlenschläger*. At denne var en af tidens berømteste digtere afficerede aldeles ikke Repp. Han kunne ikke fordra-ge ham og yppede flere gange kiv med ham. "Han maaler mig ligesom en Strikkepige maaler en Strømpe" skriver Repp i et brev til Kamma, og det kan hans stolthed ikke bæ-re. "Jeg har bestandig hørt Tale om denne Stolthed i mig siden jeg var 4 Aar", står der i samme brev fra 21 maj 1819. Og han fortsætter: "det [er] Sandhed, at jeg med meget me-

re Fornøjelse taler got om dem jeg elsker end ondt om dem jeg ikke elsker, men det nægter jeg ingenlunde, at de førstes Antal er forfærdelig meget ringere end de sidstes." Oehlenschläger, der var et venligt menneske, tog nu ikke Repp hans modvilje fortrydeligt, tværtimod. Da han vidste at Repp som en fattig student trængte til penge, engagerede han ham til at hjælpe sig med latin, som Oehlenschläger fik brug for som professor ved universitetet, og i sine erindringer omtaler han Repp venligt for hans hjælp. Endog Rahbek selv blev genstand for Repps angreb. Repp sender ham i januar 1819 et uforskammet brev, som han dog ydmygt beder om tilgivelse for i et brev til Kamma, og heller ikke Rahbek har båret nag. Han skriver i et brev fra 1822, da Repp var i England, at han "høiligen savner" ham.

En hændelse der nok har været afslutningen på Repps forbindelse med Bakkehuset beskriver J.M.Thiele i sine ovennævnte erindringer. Efter at have fortalt om Kamma Rahbeks sygelige skræk for alt hvad der lignede ligkister, skriver han: "En silde Efteraarsaften sad jeg nu med hende og ventede paa Rahbek, som skulde komme fra Byen. Da indfandt den omtalte Herre [Repp] sig til hendes ubehagelige Overraskelse i en exalteret Tilstand, der i det næste Øieblik gjorde ham utilregnelig. Det var, som om han kom, blot for at sætte hendes Nerver i Oprør eller at prøve paa, til hvilken Grad han kunde drive det i denne Retning. Efter en lang og uforstaaelig Fortælling, der gik ud paa at skildre hans Misfornøjelse med Livet, kom han til det Resultat, at han ville have en Ende paa det hele, at han allerede havde ladet sig gjøre en "Liigkiste", som han nu førte med sig paa en Vogn, der holdt udenfor osv.

Dette var nu allerede mere end nok for fru Rahbek, men hvad blev det ikke. Hidtil havde jeg moret mig over ham, men da han fremtog en lille Flaske, hvis Etiquette skulde bevise, at det var Gift, og han yttrede Tilbøielighed til straks at tage disse Draaber, fandt jeg denne Comedie mindre underholdende og aldeles ikke passende. Da Rahbek ikke var tilstede, følte jeg mig berettiget til at indskride, saa meget mere, som Fruen var midt i sine Kramper. I en myndig Tale anmodede jeg ham derfor om at holde inde med disse Drengestreger, og det lykkedes mig paa denne Maade at fjerne ham.

Jeg hørte eller saae ham aldrig siden, og det var mig af liden Interesse, da man senere fortalte, at han var gaaet til den anden Verden - til Amerika."

Det sidste maa være en fejlhuskning for hans flytning til Edinburgh.

Med denne forholdt det sig således: Repps mange interesser og vidtstrakte studier og anden virksomhed gjorde, at han ikke fik taget sig sammen til at få en eksamen, men stødet hertil kom 1825 med muligheden for at få stillingen som bibliotekar ved *The Advocates' Library* i Edinburgh. Den var først blevet tilbudt Rasmus Rask, der ifølge traditionen afslog tilbuddet med de berømte ord, som står indhugget på hans gravsten på Assistens kirkegård: "Sit Fædreland skylder man alt hvad man kan udrette". (De er dog måske ikke sagt af Rask, se herom "Bevingede Ord"). Det er vidnesbyrd om den respekt Repp nød, at han blev bragt i forslag i stedet for en kapacitet som Rask, medmindre, som N.M.Petersen spidst foreslår i sit udkast til Rasks biografi, at "vedkommende kun blev foreslået fordi man ønskede at slippe af med ham herhjemme". Forøvrigt var der allere-

de før hans ansættelse nogen betænkelighed fra The Advocates' Library's side mht. Repps temperament. Førstebibliotekaren *David Irving* skriver i et brev til Rask: "Some doubts have been suggested with regard to the smoothness of his temper," men skriver dog sidst i brevet: "I have no doubt that his good sense, and his increasing knowledge of human life, will enable him to regulate his conduct in a situation where so much will depend upon himself," en bemærkning som han bittert må have fortrudt sidenhen.

Repp indstillede sig til magisterkonferens som følge af tilbuddet, og udarbejdede en afhandling om sprogets oprindelse: "De sermone tentamen", der som led i eksamen skulle forsvares ved en (latinsk) disputation. De indledende prøver bestod han med glans, afhandlingen blev antaget med rosende bemærkninger, men forsvaret blev en katastrofe, igen foranlediget af hans vanskelighed ved at styre sit temperament.

Forsvaret fandt sted d. 6 februar 1826 kl. 10-3 i "Regenskirken", den nuværende festsal. Den første opponent var professoren i latin Børge Thorlacius, som var af islandsk oprindelse og venlig stemt over for Repp. Han bemærkede hos Repp under forsvaret nogen "Forstemthed og Distraction", som han tilskrev det rygte der havde cirkuleret om at Repp kunne vente en meget hård behandling ved disputationen, hvilket også viste sig at slå til. Efter to indlæg ex auditorio fra en student og en kandidat, tog behandlingen fat. Den tredje opponent ex auditorio var en professor i klassisk filologi, C.F.Petersen, og Repp lod sig provokere af hans indlæg, der mest handlede om formelle ting, såsom den typografiske udformning af titelbladet, til at håne ham og le ad ham. Dette fik den sidste officielle opponent, en teologisk professor ved navn Jens Møller, til at sende bud efter universitetets rektor, H.C.Ørsted, for at han kunne irettesætte Repp, hvilket han gjorde med ordene "absit risus, absit scurrilitas" (bort med latter, bort med naragtighed). Jens Møller havde af en eller anden grund et horn i siden på Repp, nogle kilder siger, at det var fordi Repp havde afsløret Møllers halvbroders teologiske disputats som et falsum, men det er usikkert. I hvert tilfælde havde Møller sat sig for at hævne sig og havde fået sig udpeget som officiel opponent, selvom han ikke var fra det filosofiske fakultet. Han tog nu fat og hegledede afhandlingen sådan igennem at Repp blev rasende og synes fuldstændig at have mistet besindelsen i sine hånlige gensvar. En senere kilde hævder at Ørsted afbrød handlingen med ordene: "Descende ex cathedra scurra" (gå ned fra katederet, nar), men det har jeg ikke fundet bekræftet i dokumenterne i universitetets arkiv. En anden kilde fortæller at Repp fik at vide at han trængte til at lære belevenhed af de danske inden han nu flyttede til Storbritanien, og det har næppe stemt ham blidere. Under alle omstændigheder blev resultatet at man nægtede at tildele Repp magistergraden på grund af usømmelig optræden under forsvaret, og denne velnok unødige strengthed skulle kaste en skygge over resten af hans liv og var sikkert en væsentlig årsag til "hans ufordulgte Ringeagt for alt Dansk" som Boye nævner.

Efter at have forlovet sig rejste Repp i begyndelsen af marts til Edinburgh for at tiltræde stillingen. Også i Edinburgh blev der vanskeligheder fra første færd. Ingen af de danske kilder nævner noget om hvordan det gik ham i Edinburgh, så det følgende er hovedsagelig hentet fra Wawns bog, der er baseret på papirer som opbevares der.

Kort tid efter at han var kommet ville han gerne have orlov for at tage til København, hvor han skulle giftes, men i stedet for blot at anføre denne grund fandt han på mystiske forklaringer til David Irving, om at hans far, som var blevet 60, ifølge islandsk lov så skulle skifte med sine børn, og at dette skulle finde sted i København. Da denne forklaring ikke blev godtaget, fandt han på en ny: nemlig at han på grund af en uberettiget anklage om at han var forfalden til en hæsleg last var nødt til at søge foretræde for kongen for at rense sit navn. I øvrigt viste det sig at han allerede havde opnået orlov fra bibliotekets bestyrelse. Kort tid efter forsøgte han at få Irving til at foranledige at hans magisterafhandling blev anmeldt i "The Edinburgh Review", og at den fik en positiv anmeldelse! Men den alvorligste episode opstod på grund af nogle kritiske bemærkninger som Irving havde ladet falde om en anden ansat ved biblioteket, en Mr Haig, og som Repp havde ladet gå videre til denne. Da Repp i Haigs nærværelse blev spurgt af Irving om han (Irving) havde sagt at Haig "was a dangerous man" svarede han straks ja, hvilket jo ikke var særlig diplomatisk. Det skaffede ham Irvings uudryddelige fjendskab og lagde grunden til en krig der varede til Repps sluttelige afskedigelse. Et tegn på hvordan situationen var ses deraf at Repp allerede i 1828 så sig nødsaget til at offentliggøre et forsværsskrift, hvori ovennævnte episode er omtalt, og hvor han påberåber sig sin sandhedskærlighed som forsvar, hvilket unægtelig lyder ejendommeligt i lyset af de andre ovennævnte episoder.

Nu rasede krigen mellem de to for fuld udblæsning, og man ser tegn på dens rasen i Rasks breve til Repp i denne periode. Allerede i et brev d. 1 jan. 1829 beklager Rask hans vanskeligheder, 20 jan. 1829 råder Rask ham til at søge til Tyskland eller Sverige (og karakteristisk fremgår det af et brev fra 14 sept. 1829 at Repp har skældt Rask ud for ikke at have takket for en gælisk ordbog Rask har modtaget). Repp forsømte sine pligter og gjorde åbenbart som det passede ham, foruden at han optrådte uforskammet. Ved én lejlighed efterlod han på Irvings skrivebord et vers fra Orlando Furioso, som kun kunne læses som en grov fornærmelse mod Irving, men benægtede at det var ham der havde gjort det.

I et brev, der findes i Det kgl. Bibliotek, fra 11. jan. 1832 til biskop P.E.Müller, som oprindelig havde anbefalet Repp til stillingen, klager Irving sin nød: "The Icelander whom you sent to the Advocates Library has proved a very unpleasant and a very unprofitable connexion; nor have I ever been acquainted with any person more destitute of gratitude and truth...Whatever may be his aptitude for learning the grammar of different languages, he is in all common affairs as stupid as a mule, and so utterly intractable that there is no method of rendering him serviceable. He obtained permission to absent himself from the Library for a few days, in order to translate a few pages of Danish into English; and on this occasion he did not return to his employment till after an interval of six months. His average attendance for the next six months amounted to about two hours a day, whereas my ordinary attendance extends to six hours. Nothing but a feeling of pity for his condition could so long have retained him in a situation where he is worse than useless."

Repps opførsel skyldtes utilfredshed med arbejdsbetingelserne, idet han fandt at det

arbejde han blev sat til var under hans værdighed, og det redegør han for i et længere forsvarsskrift som han offentliggjorde i 1834 som svar på en lang række kritiske udsagn. Mærkeligt nok er dette forsvarsskrift det eneste materiale der findes her i landet om Repps tid i Edinburgh (bortset fra Irvings brev, som jo var privat), og Det kgl. Biblioteks eksemplar stammer fra læseselskabet "Athenæum", som Repp havde tilknytning til. Man kunne fristes til at tro at Repp har en finger med i spillet her. Hvorom alting er, trods positive udsagn fra mange folk i Edinburgh (en omtaler endog hans "mildness of manner") endte affæren dog til sidst med Repps afskedigelse, og han måtte i 1837 rejse tilbage til København.

Repp følte sig uretfærdigt behandlet og vendte tilbage som en bitter mand. Han synes i det hele taget i sin selvcentrerethed ikke at have været i stand til at opfatte, hvordan hans optræden virkede på andre. Det er karakteristisk at han i sin ansøgning om stillingen i Edinburgh bruger ovennævnte Ker som reference, skønt han groft havde fornærmet Ker i et brev, og denne nægtede da også at give ham noget skudsmål overhovedet, hvad der nær havde kostet Repp stillingen.

I København fortsatte han sit rastløse virke, også for at skaffe sig et levebrød: hans økonomiske forhold var hele hans liv elendige. Han blev i et halvt år redaktør af avisen "Dagen", fik bevilling som translatør i engelsk og tysk 1839 (med karakteristisk selvbevidsthed søgte han bevilling både i engelsk, islandsk, tysk, fransk, italiensk, spansk, græsk, latin, svensk og portugisisk), underviste i engelsk, udgav bøger til brug for engelskundervisningen (herunder ordbogen) og skrev flittigt i aviser og tidsskrifter om de mest forskelligartede emner: om en forbedring af Københavns vandforsyning, som også var i en jammerlig forfatning, om indførelse af silkeorme og plantning af morbærtræer til deres føde her i landet. Han var i den forbindelse med til at stifte et selskab til fremme af silkeavl, men ragede, ikke uventet, uklar med bestyrelsen. Han deltog i debatten om den nye forfatning der skulle indføres efter afskaffelsen af enevælden i 1848, i hvilken anledning han udsendte et forfatningsudkast som indeholdt et forslag om indførelse af et overhus med adelige medlemmer, nok ikke en særlig realistisk ide, og udgav i et par år sit eget tidsskrift "Tiden" til støtte for ideen.

Men mest opsigtsvækkende var dog hans angreb på Ørsted. For at skaffe sig et fast udkomme havde Repp 1838 og 1839 søgt om at blive lektor i engelsk ved Universitetet, som ikke havde haft nogen lærer i dette fag siden 1834 (hvor den første universitetslærer i faget, Thomas Christopher Brun var død i en alder af 83). Det var en stilling hvortil han ganske utvivlsomt var særdeles kvalificeret, men man synes, måske forståeligt nok, at have været betænkelig ved at ansætte ham og få ham som kollega. Han fik ikke noget svar på sine ansøgninger og bad da i et brev 1842 om at få aktindsigt i disputatssagen fra 1826, måske fordi han mente den havde noget med det at gøre. Det blev ham nægtet, og han fik kun en kopi af den officielle afgørelse fra "Directionen for Universitetet og de lærde Skoler". Ørsted var ironisk nok en af dem der stemte for at han skulle have den fulde aktindsigt, men det kunne Repp jo ikke vide.

I november 1844 indledte Repp så en række anonyme artikler i avisen "Kjøbenhavns-

posten" med angreb på en elementær lærebog i fysik som Ørsted havde udsendt til brug for studerende ved Polyteknisk Læreanstalt. Det er svært for en ikke-fagmand at vurdere vægten af kritikken, så jeg må henholde mig til fagfolks bedømmelse, og ifølge en omtale af sagen i en udgave af Ørsteds naturvidenskabelige skrifter fra 1920 var "Angrebets reelle Berettigelse ..minimal." Repp havde altså lidt ret i sin kritik, men blæser den op til en hellig moralsk forargelse på videnskabens vegne. Hans angivelige begrundelse for angrebene fremgår af følgende citat fra en af de første artikler (Kjøbenhavnsposten 30. November 1844): "Vi vide vel, at Nogle mene, at Pietet burde afholde En fra at udpege Feil hos en saa stor og fortjent Mand som Conferentsraad Ørsted, og at man af Pietets-Hensyn hellere burde lade det Falske staae uimodsagt, fordi han havde sagt det og lært det. De, som mene dette, ere saare ilde underrettede om Pieteten: Der gives mange høiere Arter af Pietet end den personlige: Pietet for Sandhed og Videnskab er f. Ex. af en langt høiere og ædlere Art; dette havde de Gamle indseet, som sagde: "Vi holde af Plato, vi holde af Aristoteles, men vi holde dog mere af Sandheden." Dette Princip ville vi [i] alle Tilfælde følge og vi skrive ogsaa egenlig mest for dem - de være nu mange eller faa - som holde mere af Sandheden end af Conferentsraad Ørsted."

Angrebene på de ifølge ovenstående minimale fejl blev vildere og vildere. De får form af åbne breve i en højtsvungen stil til Rector Magnificus ved universitetet, hvori det erklæres at sagen er en skandale for universitetet, ja for hele landet, og hvori der trues med at sagen vil blive berettet til Videnskabernes Akademi i Paris, som Ørsted var medlem af, så han ville blive udstødt heraf.

Angrebene varede ved til august 1845, hvorpå Repp opsøgte Ørsted, afslørede sig som manden bag artiklerne og ifølge Ørsteds egen senere beskrivelse af mødet: "gjorde mig det smukke Tilbud, at ville ophøre med sine Angreb på mig, naar jeg ville skaffe ham Magistergraden, til hvis Erhvervelse han for en Snes Aar siden har disputeret, men blev afvist formedelst Forsvarets Beskaffenhed. Jeg erklærede ham strax Umuligheden heraf; men føiede dertil nogle Ord, som skulde vise, at han kun ved falske Forestillinger var bleven opægget til det Had han havde vist mod mig, og at han, som efter Naturens Lov maa overleve mig i mange Aar, let vil faae hyppig Anledning til at fortryde de Fornærmelser, hvormed han forsøgte at forstyrre den Ro, jeg nu i min Alderdom [Ørsted var på det tidspunkt 68] behøvede til at fuldende adskillige længe forberedte videnskabelige Arbejder. Man vil maaskee smile ved den godmodige Indbildning at virke paa Hr. Repp ved Grunde; men dengang syntes mine Ord at virke. Han kom dog endnu engang tilbage paa sin Magistergrad, hvortil jeg atter gav det Svar, at det, han forlangte, var umuligt, men at han muligen ved et eller andet lærd linguistisk Arbejde kunde naae det. Ved sin Bortgang erklærede han dog, at ville indgive en Ansøgning om at faae den attraaede Grad, i hvilken Anledning jeg endnu henholdt mig til mit forrige Svar." Det var altså, trods Repps ovenfor citerede smukke ord om pietet, ikke så meget videnskabens sandhed som Repps erhvervelse af magistergraden angrebene skulle fremme. Hvad den ansøgning Repp nævner angår, indsendte han d. 1. oktober 1845 en sådan til universitet. Han gav heri sin egen redegørelse for hele sagen og bad om at beslutningen fra 1826 måtte blive ændret. D. 4. november blev ansøgningen afvist uden nærmere begrundelse.

Det hører måske med i billedet at han i december 1844 atter havde indsendt ansøgning om lektoratet i engelsk ved universitetet.

Sagen fik yderligere et efterspil. D. 31. juli 1847 sender Repp et brev til Ørsted, hvori han minder denne om sit besøg "for over et Aar siden", og om det tilbud Repp havde gjort Ørsted, at han "under visse Betingelser vilde stille Debatten i Bero for det Første ... formedelst de betydelige Fordringer, jeg gjør til mig selv i Henseende til Overbærenhed, Langmodighed, Humanitet, Oprigtighed og Trofasthed i ethvertsomhelst literairt Anliggende." Men da Ørsted ikke har opfyldt Repps betingelser fra den gang, dvs. sørget for at han fik magistergraden, føler han sig "mægtig tilskyndet" til at genoptage "Debatten" af tre grunde: "1) den danske Naturlæres Renselse fra...deri herskende Feil; 2) Hævdelsen af mine egne Rettigheder og 3) Vedligeholdelsen af D[eres] H[øivelbaarenhed]s Sindsro i Deres tiltagende Alder." Det er altså under alle de smukke ord en regulær trussel, og den ledsages så af en redegørelse for den nytte han mener Ørsted har haft af påvisningen af fejlene (som var blevet rettet i senere udgaver), og insinuationer om at Ørsted jo før har støttet nogle der mindre fortjente det end Repp (Ørsted havde været beskyldt for nepotisme). Brevet fortsætter: "D.H. veed meget vel, at jeg har for mange Aar siden præsteret Alt, hvad der her er udforderligt for Erhvervelsen af en academisk Grad, og at jeg har gjort dette paa en langt forsvarligere og sømmeligere Maade, end visse andre Personer, hvis Promotion D.H. kraftigen har befordret og fremskyndet [man aner en insinuation].... Det er altsaa mit ærbødige og tillige fuldkommen venskabelige (!) Forslag til D.H.: Hæv denne Skandale". Han opfordrer Ørsted til at bruge "den Magt og Indflydelse, som De saa ofte har gjort gjældende i Sager, som vare langt mindre klare end denne [igen insinuationen]." Men nu vil han ikke være tilfreds med en magistergrad, prisen er steget: "Den høieste Grad, som Deres Facultet har at give, tilkommer mig med Rette", nu vil han have en doktorgrad, og opregner herefter de grunde Ørsted kunne anføre for at give ham den, bl.a. hans indlæg i pressen om Københavns vandforsyning og om silkeavl!

Ørsted offentliggjorde straks brevet i et indlæg i bladet "Fædrelandet", hvor han redegør for forhistorien som jeg har citeret ovenfor, og afviser beskyldningerne for nepotisme. Han mener i øvrigt at Repp næppe selv har de matematiske kundskaber som angrebene vidner om, men at han er ordfører for "saadanne Matematikere, som mene, at Physiken kun bør behandles matematisk", og redegør herefter for sit syn på forholdet mellem matematik og fysik, et syn som i dag er helt forældet. Efter nogle bemærkninger til andre punkter i Repps brev udtrykker han håbet om nu at kunne "overlade mit videre Forsvar til mine oplyste Landsmænd."

Mere kom der ikke ud af den sag, og sit formål opnåede Repp ikke ved dette forsøg, der vel kun kan forklares som en fortvivlelsens udvej for at få det som han betragtede som sin ret. Men hans brev efterlader unægtelig et forstemmende indtryk ved sin tone, der er en blanding af sleskhed, insinuationer, selvtro og trusler. Og Repp var heller ikke den der hverken fortrød eller måske var lidt flov over denne afsløring. 5 år senere, i 1852, gav han i forordet til en udgave af engelske digte til undervisningsbrug: "Udsøgte engelske Digte for Damer, for Skolen og for Studerende" en redegørelse for

hvad han selv anså som sine vigtigste fortjenester i sit virke, og her anfører han at han havde bidraget til at hæve "en paafaldende Uoverensstemmelse mellem engelsk Videnskab og Dansk Videnskab... ved at stille den i et klart Lys. Uanseet hvad Tak, jeg fik derfor, har dette siden viist sig at være et meget fortjenstligt Arbeide" (eftersom fejlene som sagt var rettet i den sidste udgave af lærebogen).

Repp fortsatte med sin utrættelige virksomhed, med forslag til en dansk forfatning i 1848 som nævnt ovenfor, med oversættelser, bl.a. af den engelske Book of Common Prayer, indlæg i pressen om medicinske, tekniske og historiske emner, og han deltog også med iver i islandsk politik. Hans håb om en fast ansættelse ved universitetet forsvandt efter udnævnelsen i 1851 af en englænder, George Stephens, til stillingen, som Repp igen havde søgt om samme år uden tanke på at hans angreb på Ørsted næppe har øget universitetets lyst til at ansætte ham; hans angreb var jo efter hans mening fuldt berettiget og havde oven i købet gavnet Ørsted.

Ifølge Wawns beskrivelse var hans sidste tid trist og præget af stor fattigdom. Han havde kun luvslidt tøj at gå i, så han måtte sige nej til middagsinvitationer hos regeringsembedsmænd som kunne have hjulpet ham. Lånekontoret og gode venners hjælp var til sidst familiens eneste mulighed for at overleve. Det hører også med i billedet af Repp at han havde gode venner, hvis venskab han beholdt, og vi får ligeledes hos Wawn et glimt af ham som familiefader, streng, men med omsorg og bekymring for sine fire børn.

Repp døde d. 5 dec. 1857, og det er måske passende som sammenfatning at citere slutningen af "Dagbladets" nekrolog over ham: "Repp var i mange Henseender begavet, besad mangesidige kundskaber, en sjelden Hukommelse og en skarp Opfattelsessevne; men disse gode Egenskaber opveiedes ved en ganske mærkelig Eensidighed og en høj Grad af Stridbarhed, der bevarede sig indtil det Sidste, saa at en af hans allerseneste offentlige Udtalelser endog indviklede ham i en Injurieproces med en Landsmand" (digteren Grimur Thomsen). Han var, som Wawn skriver "one of the most remarkable and eccentric scholarly figures ... in the first half of the nineteenth century", men mens hans ideer på andre områder ikke har overlevet hans tid, markerer ordbogen som nævnt et vendepunkt i den dansk-engelske leksikografis historie, og for den bevarer jeg ham i venlig erindring.

(Ovenstående er trykt i foredragets form og er derfor uden noter. En dokumenteret fremstilling vil fremkomme senere i en anden sammenhæng.)

De vigtigste kilder

Historien om **Webster** er hentet fra **Thomas Pyles: Words and Ways of American English**, London 1954, p. 80.

Med hensyn til **Repp** er de kilderne følgende:

Dansk biografisk Leksikon ved C.F.Bricka 1887-1905 og 2. udgave ved Svend Dahl og Povl Engelstoft 1933- 44 har artikler om Repp, men han figurerer som nævnt ikke i 3. udgave.

Th.H.Erslaw: Almindeligt forfatterleksikon... 1843- 53, med supplement 1858-68, har

en fyldig fortegnelse over Repps værker, såvel bøger som bidrag til aviser og tidsskrifter. Den mest omfattende fremstilling af Repps liv og værker er:

Andrew Wawn: *The Anglo Man. Thorleifur Repp, Philology and Nineteenth Century Britain*. Reykjavík 1991. Tyngdepunktet i bogen er tiden i Edinburgh, og Repps forhold til Bakkehuset samt affæren med Ørsted er ikke nævnt. Repps værker er udførligt omtalt eller refereret.

Forholdet til Bakkehuset er der gjort grundigt rede for i **P.H.B[oye]:** *Om K.M.Rahbeks Brevveksling*, København 1887, s. 149-162. En hel del breve fra Repp til Kamma og Knud Lyne Rahbek findes på Det kgl. Bibliotek. P.H.Boye citerer også **J.M.Thiele:** *Erindringer fra Bakkehuset*, København 1869, for bemærkningen om Lord Repp, men den refererede historie om Repps sidste besøg står først i Thieles: *Af mit Livs Aarbøger*, udgivet af Carl Dumreicher i 1907. **Oehlenschläger** nævner Repp i *Erindringer* bd. IV, s.33-34, 1851, og **Chr. Winthers** brev står i *Breve fra og til Chr. Winther* ved Morten Borup, bd.I, s. 10, 1974.

Oplysningerne om disputatsen stammer fra papirer i Rigsarkivet: dels fra Københavns Universitets arkiv, dels fra "Directionen for Universitetet og de lærde Skoler" hvortil indberetningerne fra de to officielle opponenter Jens Møller og Børge Thorlacius om disputationens forløb blev sendt. De indeholder intet om at handlingen skulle være blevet afbrudt af Ørsted. Den version står i en artikel om Repp af **Páll Eggert Olason**, i tidsskriftet *Skírnir*, Reykjavík 1916. Bemærkningen om at Repp skulle lære belevenhed af danskerne er hentet fra **Magnus Stephensens** dagbog fra et besøg i København 1826, trykt i *Personalthistorisk Tidsskrift* 8 Rk.III, 1924, men hans skildring af forløbet er i øvrigt ikke helt pålidelig. **Børge Thorlacius** blev så rasende over den behandling Repp fik at han skrev et brev til konsistorium, hvori han gav en fremstilling af sin opfattelse af sagen, bl.a. med et voldsomt angreb på Jens Møller, som han betragtede som inhabil, og bad om at få den lagt ad acta, men brevet blev dog ikke afsendt og findes nu på Islands nationalbibliotek Landsbókasafn Islands. Det citeres af Páll Eggert Olason. **N.M. Petersens** udsagn er citeret fra **Paul Diderichsen:** *Rasmus Rask og den grammatiske Tradition*, 1960, p. 218.

Skildringen af Edinburghtiden bygger hovedsagelig på Wawns bog. Den nævnte pamflet af Repp på Det kgl. Bibliotek har den langstrakte titel: *Statement to the Faculty of Advocates by Thorl. Gudm. Repp on the Subject of the Statement by the Curators of the Library*, Edinburgh 1834. **Rasks** breve er trykt i *Breve til og fra Rasmus Rask*, I-III, 1941-1968. Brevet fra Irving til P.E.Müller findes på Det kgl. Bibliotek.

Kontroversen med **H.C.Ørsted** er omtalt i *H.C.Ørstedes Naturvidenskabelige Skrifter* ved Kirstine Meyer, bd. III, s. CLV ff., 1920. Repps angreb findes i *Kjøbenhavnsposten*, diverse numre 1844 og 1845 opregnet hos Erslew i supplementet, og Repps sidste brev til Ørsted samt dennes svar står i *Fædrelandet* for 8. Augsut 1847. Originalerne til disse to breve er i Det kgl. Bibliotek. Repps ansøgning om bevilling som translator findes i Rigsarkivet.

Når maskinen tager én på ordet

- ordbogsarbejde for maskinoversættelse

Anna Braasch

Dette indlæg beskæftiger sig med nogle aspekter af ordbogsarbejdet som er særegne for maskinoversættelse af fagsproglige tekster. (For korthedens skyld bruges i det følgende de engelske forkortelser MT/HT for hhv. maskinel og traditionel (human) oversættelse).

Center for Sprogteknologi deltager i et større forskningsprojekt, Oversættelse af Fagsproglige Tekster (OFT), der støttes af Statens Humanistiske Forskningsråd. CST bidrager med erfaringer fra maskinel oversættelse. Inden for projektets rammer sammenligner vi den viden om sproget, omverdenen og fagområdet, som en oversætter gør brug af i sit arbejde, med de oplysninger, som et MT-program skal have adgang til. Det gøres ved at udføre modelforsøg: et testkorpus, der er velafgrænset med hensyn til både emneområde (domæne) og teksttype, oversættes med et MT-program.

I modelforsøget lægges der særlig vægt på arbejdet med afgrænsning af ordforråd og nødvendige oplysningstyper. Oversættelsesresultaterne synliggør fejl og mangler i ordbogen, som derefter bliver systematisk opdateret og testet. På denne måde bliver effekten af ændringer og tilføjelser målelig. Processen gentages indtil det ønskede resultat er opnået. Pilotarbejdet afsluttes med at sammenfatte de hyppigste problemtyper, der kunne relateres til ordbogen og de informationer, der var nødvendige at få systemet til at oversætte tilfredsstillende. Vore konklusioner vil også kunne inddrages i arbejdet med hjælpemidler til traditionel oversættelse.

Korpus

Teksterne, der tilsammen udgør korpuset, stammer fra domænet automobilmekanik, der hører under fagområdet mekanik, teksttypen er instruktionsbog. Testkorpuset til oversættelse er sammensat af udvalgte afsnit fra vejledninger til forskellige biltyper. Kommunikationsvejen går fra fagfolk (bilproducent) til lægfolk (bruger), hvilket bestemmer teksternes (mellemsvære) fagsproglighedsgrad. Sprogparret er engelsk-dansk. Korpus og domæne er beskrevet i detaljer i Braasch (1992a).

Oversættelsessystemet

Projektet er baseret på EUROTRA prototypen, der er et modulært opbygget, transferbaseret MT-system (jf. Copeland et al. 1991). Det har to slags sproglige komponenter: grammatikmoduler og ordbogsmoduler, af hvilke vi her vil fokusere på ordbogsmodul for engelsk-dansk oversættelse. Dette modul består af to ordbogstyper:

- de monolingvale ordbøger: engelsk analyseordbog (bruges i tekstreceptionen) og dansk genereringsordbog (bruges ved tekstproduktionen)
- den bilingvale ordbog (egentlig en ækvivalensliste), der danner bindeledet mellem den engelske analyse- og den danske genereringsordbog.

Ordbogsmodul

Systemarkitekturen er indrettet på en sådan måde, at man ved oversættelse af fagsproglige tekster kan tilkoble én eller flere ordbogsdatabaser. Dette betyder i praksis, at man kan sammensætte ordbogsmodul af to slags komponenter:

- en stor almensproglig komponent (for hvert af de involverede sprog), der indeholder det generelle ordforråd og
- mindre komponenter (satellitordbøger), der er skræddersyede til bestemte emneområder, fx petrokemi, automobilmekanik, miljøteknik etc.

Der er væsentlige fordele ved denne metode:

- man kan vælge en hensigtsmæssig kombination af almensproglig + fagsproglig ordbog (eller evt. flere fagordbøger); disse udgør tilsammen det aktive ordbogssystem
- man kan prioritere rækkefølgen af ordbogskomponenterne og derved bestemme at et ord først skal slås op fx i den biltekniske ordbog; kun hvis ordet ikke er fundet der søges der i de andre ordbøger i overensstemmelse med den fastlagte prioriteringsrækkefølge.

Ordbogsmodul bliver altså sammensat og afstemt efter domæne og teksttype, hvilket forbedrer oversættelseskvaliteten betydeligt.

Oversættelse af fagsproglige tekster

Der stilles tre elementære krav til en oversætter:

- at *forstå* kildetekstens ord og mening
- at *kende ækvivalenterne* til tekstens ord på målsproget
- at kunne *danne korrekte og adækvate sætninger* af disse ord på målsproget.

Et maskinoversættelsessystem skal også opfylde disse krav. MT-systemet udfører oversættelsen vha. tre tilsvarende moduler:

- *analyse*, med opslag i den engelske ordbog (reception)
- *transfer*, med opslag i den engelsk-danske ækvivalensliste
- *generering*, med opslag i den danske ordbog (produktion).

Et centralt problem ved oversættelse er tekstens flertydighed. Flertydigheden kan forekomme på flere planer:

- leksikalisk betinget flertydighed, når et ord har flere oversættelser
- strukturel eller syntaktisk flertydighed fx ellipse, anafor
- kommunikativ flertydighed, når kommunikationssituationen er afgørende for valget af oversættelsen.

Disse typer kan også optræde i kombination med hinanden, hvilket komplicerer oversættelsen yderligere.

En humanoversætter kan løse disse problemer ved at læse teksten i dens helhed, se på de eventuelle illustrationer, der visualiserer det faglige indhold, slå op i forskellige ordbøger og sidst, men ikke mindst ved at indhente manglende domæneviden hos fagfolk eller fra fagleksika. Vores MT-system er ikke i stand til selv at uddrage domæneviden fra den tekst, den er i gang med at oversætte (modsat mennesket, der husker det læste!). Illustrationer kan programmet heller ikke gøre brug af. Hvis maskinen altså skal kunne løse sin opgave som oversætter, skal dens ordbog eller database indeholde en stor mængde information, som ikke er repræsenteret i traditionelle ordbøger.

Når maskinen slår op i sit ordbogsmodul, skal den have adgang til de sproglige oplysningstyper, til almenviden om omverdenen og til domæneviden inden for fagområdet. Programmet kan kun arbejde med de informationer der er anført for det pågældende opslagsord i ordbogen - *maskinen tager én på ordet*.

Krav til MT-ordbøger

Kravene til MT-systemets ordbøger kan sammenfattes i nedenstående stikord.

- Generelle krav:
 - den leksikalske beskrivelse af ord og andre opslagsenheder skal være eksplicit, entydig, udtømmende (i overensstemmelse med grammatikkens krav til oplysningstyperne) og formaliseret (svarende til systemets brugersprog).
- Specielle krav i EUROTRA MT-systemet:
 - entydiggørelse af opslagsenheden skal så vidt mulig ske i analyseordbogen (hvilket her vil sige i den engelske komponent)
 - den bilingvale ordbog baseres altid på 1:1 relation mellem kildesprogs- og målsprogsenhed.

Ordbogsindgange i monolingvale MT-ordbøger indeholder i lighed med trykte standardordbøger følgende oplysningstyper:

- opslagsenhed/ord og aktuelt homograf- hhv. tydnummer, den såkaldte reading
- ordklasse og ordklassespecifikke egenskaber (fx ved danske substantiver: køn)
- morfologiske egenskaber (fx bøjningsmønster, fordobling af slutkonsonant)
- syntaktiske egenskaber (fx restriktioner vedrørende ordets syntaktiske funktion i forhold til den pågældende ordklasses generelle egenskaber; eksempelvis hvis et adjektiv

ikke kan optræde i prædikativ stilling)

- kombinatoriske egenskaber (fx valens eller brug af støtteverbum ved verbalsubstantiver, fx : 'foretage et valg')
- semantiske træk (primært ved substantiver, fx 'concrete') og selektionsrestriktioner (dvs. krav, som opslagsordet stiller til de semantiske træk ved sine valensbundne led)

En væsentlig teknisk forskel fra standardordbøgers makrostruktur er, at i denne type MT-ordbøger kodes én ordbogsindgang (dvs. en databasepost) for hver 'reading' (jf. ovenfor) af et givet ord. Medens således i COBUILD verbet *ENGAGE* anføres med 7 betydninger + 1 henvisning i én samlet ordbogsartikel, kodes der 7 (evt.8) indgange i vores ordbase.

Ordbogsarbejdet tager udgangspunkt i de ovenfor opregnede krav. I det efterfølgende gives nogle eksempler på oversættelsesproblemer, der opstår når de kodede informationer i systemets ordbog ikke er præcise eller detaljerede nok.

Problemtyper med relation til maskinens ordbog

Nogle af de hyppigste oversættelsesmæssige problemer er forbundet med maskinens krav om entydige og udtømmende leksikalsk beskrivelse af fx

- sammensætninger
- kollokationer mm.
- flertydige ord

Sammensætninger

I fagsprog er substantiverne hyppigt sammensatte af to eller flere led. I engelsk frembyder dette et specielt problem, idet sammensatte ord ofte består af en række substantiviske elementer adskilt af blanktegn. Problemet har to vigtige aspekter i MT:

- Det er svært automatisk at afgrænse det samlede kompositum, da der ikke optræder fagemærke eller andre specielle markører og da stavemåden er ret svingende: elementer står efter hinanden med blanktegn (ordmellemlum) imellem eller de er forbundet med hinanden med bindestreg, men to eller flere elementer kan også være samskrevne som ét ord. (Problemet kan forværres af at en sammensætning skrives forskelligt i samme tekst !)
- Nogle sammensatte ord kan oversættes kompositionelt, dvs. element for element til dansk (med evt. et indsat fugetegn eller strukturelle ændringer), andre kan kun delvis eller slet ikke håndteres på denne måde. Ingen trykt almen- eller fagsproglig ordbog kan opregne sammensætninger udtømmende, først og fremmest af omfangsmæssige grunde, men også fordi sammensætning er den mest produktive måde at danne nye ord på.

For at illustrere omfanget af dette problem ved udarbejdelse af ordbøger til maskinoversættelse vises her det engelske substantiv *'wheel'* ('hjul') og et udvalg af de sammensætninger med *'wheel'* der forekommer i vore tekster.

(1) *'wheel nut'*: Andet led har to ækvivalenter, nemlig den almensproglige 'nød' (frugt) og 'møtrik', hvoraf det andet klart hører til i fagområdet mekanik. I dette tilfælde kan sammensætningen oversættes kompositionelt til *'hjulmøtrik'*.

(2) *'wheel brake'*: Andet led har kun én oversættelse, 'bremse' til dansk, der også hører til i domænet. Det engelske kompositums ækvivalent er *'hjulbremse'*, der kan produceres ved kompositionel oversættelse.

(3) *'wheel brace'*: Andet led har inden for mekanik en hel række oversættelser, fx 'støtte', 'klampe', 'borsving'. Her kan kompositionel oversættelse ikke bruges, da den danske ækvivalent hedder *'svingnøgle'*.

Samme forhold gælder, når *'wheel'* er andet led i et kompositum,

(4) *'front wheel'*; første led oversættes til for(side) og ækvivalenten er *'forhjul'*.

(5) *'replacement wheel'* er lidt mindre gennemskueligt, da den første komponent oversættes med 'erstatning', 'udskiftning', når den står alene, men sammensætningen hedder *'reservehjul'* på dansk.

(6) *'steering wheel'* kan til gengæld slet ikke håndteres kompositionelt, idet ækvivalenten er et simpelt enstavelses-substantiv: *'rat'*.

Der er desuden et stort antal flerleddede sammensætninger i vores korpus, fx *'spare wheel mounting bracket'*. Langt de fleste af disse leksikalske enheder hører til inden for domænesproget.

Kollokationer

En anden type ordkombination er fx verbale udtryk, der består af et verbum med et substantiv eller en nominalfrase som objekt. Disse danner ofte en leksikalsk-semantisk enhed, en kollokation. Et typisk eksempel fra korpuset på en domænespecifik kollokation er:

(7) *'change wheel'* der altid oversættes til *'skifte hjul'*. I dette udtryk kan 'change' ikke oversættes med 'bytte' eller 'veksle' som i mange andre sammenhæng er synonyme med 'skifte'.

De ovenfor nævnte to grundlæggende problemer rejser spørgsmålet om, hvilke sammensætninger og kollokationer (og andre flerordsenheder) der principielt skal medtages i en MT-ordbog. Da en datamatisk ordbogs omfang ikke behøver at begrænses på samme måde som en trykt ordbogs, er det hensigtsmæssigt at medtage alle sammensætninger og flerordsenheder, der ikke kan håndteres kompositionelt.

Flertydige ord - flertydige sætninger

De fleste ord har flere inhærente betydninger der tilsammen udgør ordets betydningsomfang. I en given kontekst realiseres normalt kun én af disse betydninger; hvilken én, det bestemmes af den givne begrebsmæssige relation mellem nærkontekstens elementer. I kontrastiv henseende betyder dette at ordet har flere oversættelsesmuligheder. En sammensat betydningsstruktur gengives i ordbogsartiklen i almensproglige ordbøger ved at anføre nummererede betydninger, der evt. belyses med parafraser, eksempler osv.

I det følgende vil jeg koncentrere mig om sådanne problemer i oversættelsen af vore tekster, hvor leksikalsk flertydighed bliver årsag til overgenerering (dvs. for mange oversættelsesforslag) eller fejloversættelse, hvis den nødvendige domæneviden ikke er integreret i MT-systemets ordbog.

I en del tilfælde har dette problem to aspekter:

- (A) Et engelsk ord (eller leksikalsk enhed) har flere almensproglige oversættelser
- (B) Samme ord har også en række domænespecifikke oversættelser

(A) Et godt eksempel på dette fænomen er verbet '*remove*', der hører til det centrale ordforråd i vores korpus:

- (8) Remove the wheel nuts!
- (9a) Remove the jack from the engine compartment!
- (9b) Remove the hubcap from the wheel!
- (10) Remove any corrosion on the mounting surfaces with a brush!
- (11) Remove the wheel chocks and ensure that all items ...
- (12) Remove the brackets!

Verbet oversættes på forskellig vis i disse sætninger:

- (8) '*fjerne*'; (9a) '*tage frem + fra*'; (9b) '*tage af + fra*';
- (10) '*fjerne + fra* (med); (11) '*tage væk*'; (12) '*flytte*';

afhængigt af kontekstens struktur (valensmønster) og leksikalske indhold (semantiske træk), dvs. antallet og arten af verbets argumenter. I disse tilfælde kan MT-systemet håndtere valget af ækvivalent ved at sammenholde resultatet fra den sætningsstrukturelle analyse med ordbogens oplysninger om verbets valensstruktur, der i dette tilfælde altid inkluderer akkusativobjekt og i nogle tilfælde også et præpositionsled.

(B) Nedenstående eksempler illustrerer vigtigheden af, at den systemet har adgang til domæneviden. Den er kodet i ordbogen i form af semantiske træk og selektionsrestriktioner og styrer valget af den fagsproglige oversættelse i nedenstående sætning, hvori begge indholdsord er flertydige.

- (13) '*Remove the brackets!*'

Sætningen fortolkes meget forskelligt af forskellige fagfolk, fx

'Ophæv parenteserne!' (matematiker); 'Flyt lampetterne!' (bygningselektriker); 'Tag knæpladerne af!' (sømand); 'Fjern takstklasserne!' (jurist/økonom). Med opslag i ordbogen for domænet automekanik vil oversættelsen være: '*Flyt/Fjern konsollerne!*'

En anden mulighed er, at maskinen søger i den almensproglige ordbase (svarende ca. til Kjørulff-Nielsens Engelsk-danske ordbog) ved oversættelsen af følgende korpussætning

(14) '*Remove the jack from the engine compartment!*'

Resultatet vil være (med tilfældigt valgte ækvivalenter) fx 'Ryd fyren fra førerrummet af vejen!' eller 'Fjern pengene fra maskinrummet!' sammen med hel række andre, ligeså forkerte sætninger.

Problemet er at både '*remove*' og '*jack*' har flere almensproglige oversættelser, som maskinen ikke kan skelne imellem på grundlag af de oplysninger der står i ordbogen. (Et yderligere problem er at sammensætningen '*engine compartment*' ikke findes i den almensproglige ordbog, dvs. systemet forsøger at oversætte den kompositionelt.)

Men, som det allerede er blevet omtalt i afsnittet om systemarkitektur, foretages der prioriteret ordbogsopslag (først i domæneordbogen og derefter i den almensproglige komponent) og sætningen kan dermed oversættes korrekt til '*Tag donkraften frem fra motorrummet!*'

Ordbogsarbejdet

Ordbogsarbejdet - også inden for MT - handler i høj grad om at genbruge eller udnytte tilgængelige oplysninger i eksisterende opslagsværker. Foruden at gennemgå og bearbejde det genbrugelige materiale udfører vi også korpusbaseret ordbogsarbejde. Korpusundersøgelser fokuserer på et ords nærmeste omgivelser; i vores arbejde betyder dette altid analyse inden for den aktuelle sætnings grænser.

Arbejdsmetode

Den maskinelle korpusanalyse kan gøre brug af mere eller mindre sofistikerede metoder, afhængig af, hvor store tekstmængder der skal analyseres, til hvilket formål, og hvilket program man har til rådighed. Vores korpus på ca. 50 A4 sider fylder ikke ret meget i maskinlæsbar form.

Vi anvendte først et større konkordansprogram (WordCruncher), der har gode søge- og sorteringsfaciliteter, men er noget omstændeligt i brug. Programmet er blevet benyttet i den indledende fase til at fremstille en samlet liste over alle ordformer samt deres forekomststal i korpuset. Desuden blev der fremstillet en samlet nøgleordskonkordans (KWIC). Dette materiale dannede grundlaget for det mere målorienterede arbejde, der har involveret en række manuelle og automatiske processer der beskrives nedenfor.

Hertil valgte vi at bruge programmet Excerpter, der er mere egnet til hurtige detaljeundersøgelser på grund af sin smidighed og enkle, men effektive virkemåde.

Vi har foretaget:

- selektion af leksikalsk materiale til videre behandling med hensyn til de tidligere opregnede oversættelsesmæssige aspekter (fx led i sammensætning hhv. kollokation; flertydighed mm.)
- syntaktisk analyse af nøgleordets forekomster med henblik på valensmønstre, kollokationer
- sortering af forekomsterne i valens- og kollokationstyper
- semantisk analyse af nøgleordets kontekst (primært af de valensbundne led, jf afsnittet om 'Krav til MT-ordbøger')
- leksikalsk analyse der omfatter sammensætninger, brug af synonymer, antonymer mm.
- morfologisk analyse (bøjningsformer, kongruens, stavevarianter)

Foruden at dele de observerede sproglige træk op i typer og dermed forberede dem til beskrivelse i systemets formelle sprog er der også behov for at beskrive visse grundlæggende, generelle begrebsrelationer inden for domænet (jf eksemplerne 8 til 12, 'remove').

Sådanne relationer er den såkaldte "is_a" relation (forklaring ved nærmeste overordnede begreb, fx 'wheel_nut' is_a 'fastener'; "is_part_of" relation (del relateres til helhed fx 'hubcap' is_part_of 'wheel') og "functions_as" (der relaterer genstanden til en funktion inden for domænet, fx 'jack' functions_as 'lifting_tool'). En del af disse oplysninger kan udtrages fra selve korpusteksten, resten kan findes i de tilhørende illustrationer eller indhentes fra eksterne kilder.

Disse oplysninger sættes i system: begrebernes indbyrdes relationer registreres i et hierarki (som over/under- hhv. sideordning) og til slut opstilles en taksonomi over domænet. De relationer, som taksonomien er baseret på kan udtrykkes i maskinens formelle sprog. På denne måde tilføjes en slags domæneviden til ordbogen på linje med de sproglige oplysninger. Herved kan entydiggørelse af kildeteksten optimeres og valget af oversættelsesækvivalent styres. På samme måde fungerer definitioner og forklaringer i en tosprogsordbog.

Konklusion

Den maskinelle oversættelse stiller krav om fuldstændig og systematisk beskrivelse af tekstens ordforråd. Opfyldes disse krav ikke, bliver oversættelsen mangelfuld eller forkert. De her præsenterede oversættelsesproblemer har væsentlige lighedspunkter med de vanskeligheder, en oversætter kan komme ud for, når der mangler opslagsord eller informationer i den anvendte fagordbog. Et andet problem kan være, at den anførte information er flertydig eller ikke er detaljeret nok; selv ved hjælp af kreativ tænkning kan man ikke altid finde frem til den korrekte oversættelse.

Ved at overføre relevante erfaringer fra arbejdet med maskinel oversættelse til planlægning af fagordbøger burde man kunne nå frem til en mere målrettet valg af ordforråd

og oplysningstyper. Samtidig kunne præsentationen blive mere konsistent end det er tilfældet med de ordbøger, vi har undersøgt. Man kunne med fordel integrere begrebsmæssige relationer i ordbogsartiklernes definitionsdel, i lighed med dem der bruges i MT.

Hensigten med at beskrive nogle centrale problemer i MT med relation til ordbogen har været at vise, hvorledes traditionel fagleksikografi kan drage nytte af det lingvistiske arbejde, der udføres i forbindelse med forskningsprojektet OFT.

Litteratur

- Braasch, A. (1992a.) *Valg af tekstsort - Korpus - Undersøgelsesaspekter*. In: Ark 65, Oversættelse af fagsproglige tekster. HHK. København. 1992.
- Braasch, A. (1992b). *Text based dictionary work for a domain-specific language*. In: Papers in Computational Lexicography. Budapest. 1992.
- Copeland, C., Durand, J., Krauwer, S. & Maegaard, B. (eds.) *Studies in Machine Translation and Natural Language Processing Volume 1: The Eurotra Linguistic Specifications*. Luxembourg: CEC. 1991.
- Sager, J. C. *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 1990.

Ordbøger

- Clausens tekniske ordbøger. Engelsk-dansk.
- Collins COBUILD English Language Dictionary. London & Glasgow, 1987.
- Kjærulff-Nielsen, B.: Engelsk-Dansk Ordbog. Gyldendal, København, 1985.
- Skibsted, S.: Teknisk engelsk-dansk ordbog. København, 1971.
- Skjerk, Ebbe: Bilteknisk ordbog. Engelsk-Dansk. Teknisk Forlag, 1991.
- Teknisk ordbog. Engelsk-dansk. L&H Ordbøger. København, 1991.
- 950 amerikanske-engelske automobil-fagudtryk oversat til dansk. General Motors International A/S

Programmer

- Excerpter, Version 2.0. Produceret af Norling•C Dataleksikografi, København
- WordCruncher, Version 4.23. Produceret af Brigham Young University.
- ElectronicText Corporation, Provo, Utah.

EGs ord och uttryck.

Om TNCs arbete med en korpusbaserad, deskriptiv ordlista

Anna-Lena Bucher

Bakgrund

När det i juni 1990 stod klart att delar av EGs regelverk (ca 10 000 sidor) skulle översättas till svenska i samband med EES-avtalet, inrättades en särskild myndighet för ändamålet, "Delegationen för översättning av EGs regelverk" (översättningsdelegationen). Själva översättandet fördelades på 35 olika översättare som var och en knöts till en arbetsgrupp bestående av 3 - 4 experter.

Hur terminologin skulle hanteras var tidigt en av de frågor som stod i centrum när delegationen skulle organisera arbetet. Direktiven, råden och rekommendationerna rörde så vitt skilda områden som etableringsrätt, offentlig upphandling, konsumentskydd, statistik, produktansvar, teletjänster, arbetsrätt och miljö. Det var uppenbart att dokumenten skulle innehålla en stor mängd facktermer som översättningsdelegationen inte så lätt skulle kunna samordna användningen av eftersom arbetet var fördelat på så många händer.

Det ledde till att TNC kopplades in i diskussionerna och fick i uppdrag av översättningsdelegationen att bygga en engelsk-svensk EG-termbank samt att ge råd och upplysningar åt översättarna om enstaka termer.

Termbanksbygget

Översättningsarbetet brådskade och det var ont om ordlistor som var passande för ändamålet. Varje ord eller uttryck som en översättare i onödan behöver tänka länge över, rådfråga utomstående om, slå upp i ordböcker, skapa en ny term för etc, kostar tid och pengar. Därför beslöt TNC att som första material i termbanken lägga in terminologi ur de två officiella EG-dokument som redan fanns i svensk översättning: *Romfördraget*, dvs *fördraget om upprättandet av den europeiska ekonomiska gemenskapen* och *Vitboken*, dvs *EG-kommissionens vitbok om förverkligandet av den inre marknaden*.

Vitbokens engelska och svenska versioner excerperades helt manuellt på TNC och excerpterna registrerades i TNCs termpostformat, se figur 1.

- enTE procurement policy (engelsk term)
- svTE upphandlingsprincip (svensk term)
- enKX where possible procurement policies have a major impact (engelsk kontext)

- svKX där eventuella upphandlingsprinciper får stor effekt (svensk kontext)
- UR White Paper 1985/EG-kommissionens vitbok, par 87(källa)

Figur 1

Excerperingen

För att excerperingen skulle bli så enhetlig som möjligt bestämdes vissa urvalskriterier. I första hand skulle man excerpera EG-specifika begrepp som snabbt borde få en enda accepterad, välvald term, t ex *infringing State* (stat som överträder EG-rätten). Den som excerperade skulle i andra hand välja sådant som såg "termigt" ut, dvs ord och uttryck som verkade stå för fackspecifika begrepp och som troligen kunde vara svåra att finna motsvarigheter till i andra lexikon. Så excerperades ur Vitboken t ex *awarding entity* (upphandlingsmyndighet) och *re-partitioning* (förnyad uppdelning).

Genom att fånga upp redan översatta termer inuti löpande text och göra dem sökbara var för sig, erbjöd TNC översättarna ett verktyg som kunde begränsa antalet onödiga varianter, ja, rensa i en vilt växande flora av "synonyma" termer. Ändå uttalade vi klart, efter samråd med översättningsdelegationens ledning, att de svenska termmotsvarigheternas status var sådan att de inte var bindande för det fortsatta översättningsarbetet. Eftersom de ingick i givna kontexter kunde de svenska termerna inte betraktas som generella motsvarigheter till de engelska; de var heller inte i sitt sammanhang alltid de bäst tänkbara. Det sätt som termerna ställdes samman och presenterades på skulle snarare ses som ett stöd och beslutsunderlag i översättningsarbetet än som ett facit.

Det andra EG-dokumentet som fanns i svensk översättning var Romfördraget. EG-språkens version av Romfördraget hade redan excerperats vid EG-kommissionens översättningsavdelning i Luxemburg. Dessa excerpter kunde TNC få tillgång till i maskinläsbar form. Ur det materialet valde vi att arbeta med excerpter på endast engelska och franska. De svenska motsvarigheterna fick vi själva hämta ur den översättning som fanns.

I arbetet med Romfördraget upptäckte vi att det inte var enbart rena termer som excerperats i Luxemburg utan också fraser och kollokationer, såsom *a case pending before a court or tribunal of a Member State* eller *the powers conferred by this Treaty*. De passade inte in i TNCs traditionella termpostformat utan lades tills vidare åt sidan.

I december 1990 gjordes ett första utdrag ur termbanken i form av ett kompendium som kallades Basordlista och som distribuerades till alla de 35 översättarna och arbetsgrupperna.

Hur göra med fraserna?

Men fraserna som blev över ur Romfördragets materialet - vad skulle vi göra med dem? Vi konstaterade att de ur översättningssynpunkt säkerligen var minst lika knepiga som

termerna, och därför var det viktigt att de kom med i termbanken.

För att göra det möjligt att söka efter enstaka ord i fraserna gjorde vi ett konkordansregister, se figur 2.

| | |
|-------------------|---|
| nger capacity and | market access - air transport |
| es on the capital | market and long-term commercial credits |
| internal nuclear | market and the validity of such conditions unde |
| treat the common | market as a single environment for their econom |
| ent of the common | market for agricultural products must be accomp |
| ent of the common | market for broadcasting, especially by satellit |
| creating a common | market for services, all those who provide and |
| an expanded home | market for the people and the industries of the |
| ional producers a | market for their products |
| o take account of | market forces, |
| the cross-border | market in the traditional services, |
| future internal | market initiatives |
| The information | market is also undergoing far reaching changes |
| ion of the common | market is not affected through distortions of t |
| be able freely to | market its units throughout the Community |
| ishing the common | market must be implemented |
| he placing on the | market of medical specialities |
| on the domestic | market of the purchasing country |

Figur 2

Frasposterna med konkordansregistret presenterades i en speciell ordlista som också distribuerades till översättarna och arbetsgrupperna.

En av översättarna, Adolf Dahl, har beskrivit översättningsarbetet och då sagt följande om de två dokumenten:

"På plussidan var det tre faktorer som underlättade arbetet. Först och främst hade vi tillgång till specialordlistor som utarbetades speciellt för regelverket. I december 1990 kom Basordlista och i mars 1991 Frasordlista, båda utarbetade av TNC. Detta stöd i översättningsarbetet var av stort värde. I synnerhet frasordlistan med färdiga konkordanser var ett ganska ovanligt hjälpmedel som flitigt kom till användning. Kontexter från Romfördraget och EG-kommissionens vitbok var tänkta som ett stöd men blev ofta ett formligt facit både när översättaren sökte en bra term eller ville kontrollera likformigheten. Även till synes enkla termer som *Member State* eller *construction product* kunde med hjälp av ordlistorna redas ut i bästa sämja när det blev diskussion i arbetsgruppena".

Excerperingen av de nyöversatta dokumenten

Allteftersom de nya preliminära översättningarna blev färdiga excerperades de vid TNC. Men vi utvidgade nu urvalskriterierna (efter påverkan från arbetet med Romfördraget): förutom rena termer excerperade vi också fraser, vilka kunde vara fackspråkliga idiom, kollokationer och tillfälliga ordkombinationer.

Vid en första anblick uppfattades t ex *defective product* och *exonerating circumstance* som engelska termer medan de svenska motsvarigheterna *produkt med säkerhetsbrist* och *omständighet som befriar från ansvar* verkade betydligt mindre "termiga".

Put a product into circulation och *claim pending at law* noterades som engelska fraser. Motsvarigheterna på svenska var *sätta en produkt i omlopp* och *mål som redan väckts*. Gränsen mellan fras och term var inte alltid alldeles självklar och det föranledde en hel del diskussioner.

I oktober 1991 innehöll den svenska EG-termbanken 4 500 poster vilka var hämtade ur den preliminära översättningen av EG-dokumentet. Preliminär betyder här att det översatta dokumentet inte hade slutgranskats av juristlingvister i Bryssel. EG-termbanken lades tillsammans med nio andra termbanker in på en CD-ROM-skiva, "Termdok på CD-ROM", som TNC tillsammans med Walters Lexikon publicerade i mars 1992.

Planer på att ge ut innehållet i termbanken även i bokform och anpassa det så att det kunde nå en vidare krets av användare, började då göras upp. I maj 1992 tecknade TNC ett avtal med Bra Böcker om att ge ut en engelsk-svensk EG-ordlista för bl a översättare med ca 4000 poster. Ordlistan skulle baseras på innehållet i termbanken.

Riktlinjer för ordlistearbetet

När ordlistearbetet inleddes innehöll termbanken ca 8 000 poster med termer, namn och fraser hämtade ur 400 dokument, och vi måste alltså göra ett urval. Ett antal riktlinjer drogs upp innan arbetet startade.

Ordlistan skulle innehållsmässigt omfatta områdena: ekonomi, juridik, handel, EG-administration, teknik och övrigt (t ex sådant som är så EG-specifikt eller nytt att det kunde vara mycket svårt att översätta). Dessutom skulle ordlistan som typ vara passiv (från engelska till svenska), deskriptiv och korpusbaserad.

Ordlistan är både deskriptiv och normativ

Från början var alltså tanken att ordlistan skulle bli deskriptiv. Men i dag ser vi resultatet: ordlistan är både deskriptiv och normativ.

"Sant" deskriptiv är den genom att kontextuella översättningar redovisas, ibland flera för en och samma engelska term.

"Begränsat" deskriptiv är ordlistan genom att den är baserad på officiella, språkligt EG-anpassade, slutgranskade dokument, att översättningarna är granskade av experter och att en del felaktiga eller missvisande termer för ett visst begrepp har uteslutits för att de inte i onödan skall spridas vidare.

Nedan finns exempel på oprecisa eller nyskapade termer som vi fann och som vi valde att stryka:

| <i>Engelsk term</i> | <i>Svensk översättning</i> | <i>Bör hellre vara</i> |
|-----------------------|----------------------------|------------------------|
| domestic hot water | hushållsvarmvatten | tappvarmvatten |
| employer contribution | arbetsgivarens avgift | arbetsgivaravgift |
| post mortem unit | bårhus | obduktionsavdelning |
| vocational test | anlagsundersökning | anlagstest |
| occupational disease | arbetssjukdom | yrkessjukdom |

Men ordlistan är också normativ genom att den innehåller rekommenderade översättningar av namn bl a på råd, organisationer och kommittéer samt rekommenderat skrivsätt för val av versal eller gemen begynnelsebokstav. Vissa av namnposterna är försedda med förklaringar som är baserade på en eller flera auktoritativa källor, se figur 3.

**Committee of Permanent Representatives;
COREPER; Permanent Representatives
Committee**

Ständiga representanternas kommitté
(medlemsstaternas ambassadörers kommitté som bereder ärendena inför rådets beslut; COREPER = Comité de représentants permanents)

European Coal and
Steel Community;
ECSC; CECA

Europeiska kol- och
stålgemenskapen; EKSG
(CECA = Communauté
européenne du charbon et
de l'acier)

rekommenderat skrivsätt
på svenska

rekommenderad svensk
förkortning

uttydning av
fransk förkortning

figur 3

Även om vi säger att ordlistan är deskriptiv så är vi givetvis medvetna om att även en deskriptiv ordlista med tiden får en viss normerande verkan, såsom Adolf Dahl beskriver. Men det är ändå viktigt att påpeka att posterna visar hur en viss översättare i en viss situation har valt att översätta en viss term eller fras. Den översättning som redovisas kan vara den bästa vid flera tillfällen, men man kan i en annan kontext behöva välja en annan lösning. De redovisade översättningarna är således inte alls bindande utan skull, precis som översättningarna i termbanken, utgöra en del av ett beslutsunderlag.

Både lexikografisk och terminologisk metod

När vi arbetat fram ordlistan har vi använt oss av en kombinerad lexikografisk och terminologisk metod. Den lexikografiska metoden har visat sig i att ordlistan är baserad på

en enhetlig korpus vilket också innebär att huvudkällan är skriftlig, att den huvudsakligen utarbetats alfabetiskt och att allmänspråkliga uttryck, särskilt prepositionsuttryck, tagits med.

Exempel på sådana prepositionsuttryck är *by common accord* (efter överenskommelse, gemensamt, i samförstånd), *by way of derogation from* (trots, utan hinder av), *without prejudice to* (utan att det påverkar tillämpningen av).

Terminologiskt arbete kännetecknas av att det utgår från begreppet. Detta faktum har färgat av sig i ordlistans uppställning genom att nästan alla poster utgår från ett och endast ett fackspecifikt begrepp. Även fraserna hanteras som termer och redovisas därmed i egna poster. Flerordstermer, liksom fraser och prepositionsuttryck sorteras under första ordets första bokstav. Exempelvis sorteras *by virtue of* under *by* men det finns också en s k naken hänvisning från *virtue*. *Temporary employment relationship* sorteras under *temporary* men det finns också en hänvisning under posten *employment relationship*. *Open market economy with free competition* sorteras under *open* men en hänvisning ges från *free competition* eftersom detta betraktas som ett betydelsetungt uttryck inuti uppslagsordet.

På traditionellt terminologiskt vis har vi också haft ämnesvisa genomgångar med experter som bedömt såväl urval av begrepp som termernas korrekthet. Dessutom har ordlistan varit utsänd på remiss till språkexperter med särskild erfarenhet av hur EGs regelspråk översätts till svenska.

Några särskilda stötestenar

"principle"

Det visade sig att vi hade excerperat en mängd kollokationer som inleddes med ordet "principle". Eftersom "principle" i sig inte är svårt att översätta beslöt vi oss för att ta bort ordet och bara behålla det efterföljande i kollokationen. Vissa kollokationer med "principle" behöll vi emellertid. Det gällde de fall då principen ifråga nästan var av namnkarakter, dvs en princip som antagits av EG. Vi hade att utgå från:

- principle of substitution *
- principle of supervision on a consolidated basis
- principle of the comparability of university studies between the Member States
- principle of mutual recognition
- principle of mutual trust
- principle of non-discrimination by nationality
- principle of non-discrimination *
- principle of proportionality *
- principle of reciprocity *
- principle of safety
- principle of subsidiarity *
- principle of coordination *
- principle of equal pay *

Anna-Lena Bucher

principle of equal treatment for men and women
principle of equal treatment*
principle of freedom of capital movements
principle of home control

Vi diskuterade frågan med Peter Ammundsen vid EG-kommissionens översättningsavdelning i Luxemburg. Han kunde så småningom hjälpa oss att skilja ut vilka uttryck som står för begrepp och vilka som är mer tillfälliga sammansättningar; vi behöll "principle" framför de kollokationer som markerats med asterisk ovan. I de övriga fallen behandlade vi i stället huvudordet i uttrycket som uppslagsord.

Sammanfattningsvis kan man säga att det var experter som fick avgöra vilka uttryck som var termer; för en lekman var det omöjligt.

"gemensamma marknaden"

Vi observerade också att tre termer användes på engelska för ett väldigt centralt EG-begrepp, nämligen det som på svenska oftast kallas "den gemensamma marknaden". Det var *the common market* (den gemensamma marknaden), *the internal market* (den inre marknaden) och *the single market* (den enhetliga marknaden).

Vi trodde att en av dessa termer var den som rekommenderades i första hand inom EG och att de andra två var mer beskrivande eller tillfälliga sammansättningar. Trots att vi arbetade med en deskriptiv ordlista ville vi gärna föra in en anmärkning om dessa tre uttrycks olika status, eftersom begreppet var så EG-centralt.

Men - genom våra kontakter med Luxemburg fick vi veta att det inte fanns någon central rekommendation om vilken term som i första hand borde användas. I stället förhöll det sig så här:

common market är ett nyckelbegrepp i Romfördraget och har också övertagits i Maastrichtfördraget.

internal market introduceras på allvar i Vitboken (1985) och enhetsakten (1986); man kan eventuellt hävda att *internal market* är ett underbegrepp till *common market* som skulle definieras som "gemensam marknad utan inre gränser".

single market är ett uttryck som ibland används synonymt till *internal market* som en stilistisk variant, ibland kan man också se *single internal market*.

Det visade sig alltså att det inte var så som vi som terminologer hade lätt för att tro: att det fanns en klar och entydig rekommendation om termval för ett centralt begrepp. Alla tre posterna finns med i ordlistan men utan det tänkta tillägget.

Redaktionell bearbetning

En bedömning av engelskan har utförts av Patricia Thomas, med brittisk engelska som modersmål. Hon har bedömt vilka fraser som tillhör allmänspråket och som kan förvän-

tas vara allmänt kända. Hon har också hjälpt oss att få uppslagsorden i grundform och grammatiskt korrekta, t ex om det skall vara *pay for work at time rates* eller *rate* (skall vara *rates*), *misuse of powers* eller *power* (skall vara *powers*).

Bo Svensén, som har kunskaper i både lexikografiskt och terminologiskt arbete, har värderat posterna och bedömt uppslagsordens form, såväl på engelska som svenska.

Vi var tvungna att göra en hel del redaktionell granskning. Ett av de främsta skälen var att Romfördraget under hösten 1992 kom i ny version. Dessutom hade direktiven slutligt granskats av juristlingvister i Bryssel och dessa s k slutversioner var färdiga under senhösten 1992. Detta gjorde att vi jämförde de slutliga versionerna med de preliminära och förde in eventuella ändringar.

Ändringar som genomgående införts i slutversionerna var t ex *sammanställd redovisning* i stället för preliminärversionens *koncernredovisning*, *på gruppbasis* i stället för *på koncernbasis*, *egenföretagare* i stället för *självständigt förvärvsverksam person*.

Termpostformatet i ordlistan

Vi hade med jämna mellanrum diskussioner om hur termpostformatet skulle se ut i ordlistan. Den slutliga formen bestämdes strax före typograferingen. Termpostformatet i ordlistan bygger givetvis på TNCs termpostformat i termbanken men har anpassats med tanke på användningen och innehållet, se figur 4. Ordlistan har också ett svenskt register.

| | | |
|---|-------|--------------------------------|
| public law | _____ | engelskt uppslagsord |
| offentlig rätt | _____ | svensk term |
| <i>the State, regional or local authorities and the legal persons governed by public law</i> | _____ | engelsk kontext |
| statliga, regionala eller lokala myndigheter och de juridiska personer som lyder under offentlig rätt | _____ | svensk kontext |
| [EGT L 185/5 1971, s 683] | _____ | källa |
| → <i>contract governed by public law</i> | _____ | hänvisningar till andra poster |
| → <i>fall under public law</i> | _____ | |

Figur 4

Slutord

TNC skall nu fortsätta att utöka termbanken genom att excerpiera uttryck ur de delar av EGs regelverk som är aktuella för ett eventuellt svenskt medlemskap och som just börjat översättas. Så småningom kommer vi också att komplettera termbanken med relevanta

definitioner ur TNC-ordlistor, svensk standard, andra normativa dokument såsom föreskrifter, anvisningar, allmänna råd etc. Dessutom behöver de ca 4 000 termposter i termbanken som inte ingår i ordlistan granskas och bearbetas på samma sätt som posterna i ordlistan. Det långsiktiga målet är att vi efter hand skall kunna rekommendera en av flera termer eller fraser framför de andra.

Vi har också planer på att komplettera termbanken med motsvarigheter på de nordiska språken. Ett sådant samarbete har redan inletts med TNCs systerorganisation Tekniikan Sanastokeskus (TSK) i Finland, som för närvarande fyller på termposterna med finska ekvivalenter hämtade ur de finska översättningarna av EGs regelverk.

En väl utbyggd nordisk termbank är ett viktigt verktyg i de kontakter - gemensamma eller enskilda - som Norden kommer att ha med övriga Europa i framtiden. En sådan termbank är minst lika viktig även om Sverige, Norge och Finland inte kommer att bli medlemmar i EG. Vid ett medlemskap kommer en hel del nordiska begrepp att successivt harmoniseras med de existerande EG-begreppen. Men begreppsharmonisering är ofta en långsam process och medan den pågår är det viktigt att ha kunskap om olikheterna begreppen emellan. Blir vi inte medlemmar måste vi, mer än vad vi gör i dag, tydligt dokumentera skillnaderna mellan våra begrepp och EGs.

(Föredraget framfördes på konferensen av Katja Hallberg och Lars Törnqvist, båda TNC.)

Idiom och variation

Ulla Clausén

Idiom så som de normalt registreras i gängse ordböcker är visserligen ganska sällsynta i löpande text, men de används, och då ofta i en varierad form.

De klassiska idiomerna ingår tillsammans med talesätt och jämförelser i ett s.k. idiomblock. Fraserna i detta block har en dynamik, som bl.a. yttrar sig så att de rör sig mellan kategorierna: idiomerna varierar som talesätt och jämförelser, talesätten och jämförelserna som idiom. De idiomatiska uttrycken finns inom en glidande skala från det konventionella mönsteruttrycket, som ibland kan uppfattas som banalt, till det tillfälliga och individuella uttrycket.

Gängse ordböcker har i regel förbisett att det är viktigt att exemplifiera de idiomatiska uttrycken, dvs. visa hur de kan användas. En konstruktionsordbok som ska kunna vara till verklig hjälp i skrivandet bör dels ange den variation som finns på frasnivå, dels antyda de möjligheter som finns på nivån under frasnivån, på exempelnivå.

Vid Svenska språknämndens sekretariat i Stockholm är sedan en tid tillbaka en konstruktionsordbok under arbete; det är den gamla och beprövade Svensk handordbok från 1966 som genomgår en omfattande bearbetning. Den nya ordboken, Svensk konstruktionsordbok (SKO), kommer att bli mer specialiserad än sin föregångare: en renodlad konstruktionsordbok. Materialet innefattar dels produktiva konstruktioner, dels halvfasta och fasta fraser. Bland de fasta fraserna finns idiomerna (se figur nedan).



Figur 1

I vår kategorisering av de fasta fraserna finns tre grupper: 1) lexfraser 2) jämförelse-idiom – talesätt 3) klichéer– sociala fraser. Bland dessa utgör idiom en relativt liten grupp bestående av s.k. klassiska eller egentliga idiom av typen *kasta pärlor för svin* ('ödsla ngt värdefullt på ngn som ej kan värdesätta det'). Det är fråga om metaforiska uttryck som är mer eller mindre fasta till sin form. Utmärkande för idiomerna är också en dynamik och en modalitet, som lexikografen bör ta hänsyn till vid presentationen av idiomerna i en ordbok.

Det har framhållits att idiom är relativt sällsynta i löpande text (Moon 1992 s. 493). Ändå ger gängse ordböcker relativt stort utrymme åt dem. Frågan är om detta kan vara befogat? I den form idiomerna registreras är de visserligen inte så vanliga, men de finns i andra former. Idiomerna behöver upptäckas på nytt.

Vissa delar av ett idiom kan ibland varieras, andra delar aldrig. Hur lexikografen ska ställa sig till variation och praktiskt lösa frågan i ordboksarbetet kan vara värt att diskutera. Variation av idiom tas visserligen upp i några handböcker, men huvudsakligen från textlingvistisk synpunkt (jfr Burger 1982).

Gängse ordböcker är ofta sparsamma med uppgifter om när variation av idiom är möjlig och vad som i så fall kan varieras. Om en ordbok ska ha med sådana uppgifter måste det emellertid framgå vad som är ett uttrycks normalform och vad som är variation. I den nya ordboken kommer mönsterfraser att få en formaliserad form som lyfts fram i typografin genom en särskild stil (fet kursiv); de variationer som tillhör normalformen får samma stil, med varianten placerad efter ett alternativtecken eller inom klammer. Språkexempel ges i brödstil. Med hjälp av typografin ska olika nivåer i ordboken klart framgå: frasnivå, exempelnivå och kommentarnivå. I artikeln "skinn" finns exempel på dessa nivåer: frasen "vara [bara] skinn och ben" i fet kursiv, exemplet "hon var bara skinn och ben de sista åren" i brödstil och kommentaren "som karakteristik av mycket mager person" i rak petit (se figur nedan).

skinn subst., oftabildl. *ömsa skinn* det fina varuhuset har ömsat ~ och satsar på enklare varor; bildl. *rädda/klara sitt [eget] skinn* (för egen del undgå förlust el. obehag) han dolde sanningen för att rädda sitt eget politiska ~; bildl. *vara rädd om sitt eget skinn/det egna skinet* (vara feg) jfr *rädda, rädd*; bildl. *med skinet i behåll* (utan förlust) han kom ifrån den trassliga affären med ~et i behåll; *in på bara skinet* (in på barakroppen) han kom hem efter äskskuren vänt in på bara ~et (genoblöt); turisterna blev plundrade in på bara ~et

I jämförelse, vard., ngt nedlåtande om stressad person [*fara/flänga/rusa omkring*] som *ett torrt skinn* (rusa runt för att hinna med mycket) hon far omkring som ett torrt ~ för att hinna med alla extrajobb, äv.

Figur 2

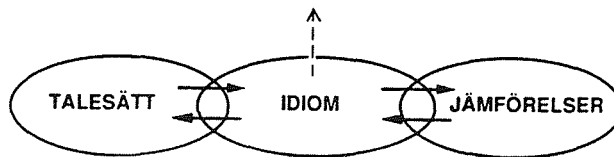
fara omkring som en skällad rätta se rätta; som karakteristik av mycket mager person *vara [bara] skinn och ben* hon var bara ~ och ben de sista åren; vard., vanl. positivt omdöme om viljestark person *ha skinn på näsan* måste ha rejält med ~ på näsan för att våga gå emot ledningen; vard., om uppträdande *hålla sig i skinet* (hålla sig istyr), *krypa ur skinet* (bli utom sig) hon var färdig att krypa ur ~et av ilska (hon var mycket orolig); talesätt: [*man ska inte*] *sälja skinet innan björnen är skjuten* (man bör inte räkna med ngt som man ännu inte har); *arga katter får rivet skinn se katt*

I opersonliga verbuttryck, för att uttrycka smärta *det svider i skinet* det sved ännu i ~et efter örfilen; bildl., vid ekonomisk förlust det sved ordentligt i ~et när jag skulle betala räkningen; *det kryper i skinet* (det känns olustigt) det kröp i ~et på henne när hon hörde hans namn nämnas

Typiskt för idiom är att i dem ofta ingår attityder eller konnotationer som är svåra att få fram i en enkel definition. Vi vill i kommentarer göra användaren uppmärksam på dessa. I artikeln "skinn" finns t.ex. idiomet *ha skinn på näsan* som inleds med kommentaren "vanl. positivt omdöme om viljestark person". Att ge en heltäckande definition av ett idiom är ofta mycket svårt. Först när de tre nivåerna, frasnivå, exempelnivå och kommentarnivå har utnyttjats för att ge all tänkbar information kan idiomets sägas vara uttömmande behandlat.

Variation inom idiomblocket

Bland de fasta fraserna finns den grupp om tre kategorier, som förutom idiom innefattar talesätt och jämförelser. Talesätten är hela satser av typen *man ska ta seden dit man kommer*, som ofta har innebörden av en allmän levnadsregel. Jämförelserna är uttryck av typen *fara / flänga omkring som ett torrt skinn*. Dessa tre kategorier ingår i ett block, jag vill kalla det "idiomblock", där idiomerna utgör centrum, och där talesätt och jämförelser finns i ytterkanterna. Många talesätt är på väg att närma sig ett idiom, men ändå kvarstår innebörden av ett moraliskt påpekande eller levnadsvisdom. Idiomerna kan i sin tur närma sig de jämförande uttrycken, genom att ett "liksom" e.d. tillfogas. Här finns länken från jämförelserna. En jämförelse kan också närma sig ett idiom genom att det jämförande *som* stryks. Slutligen kan ett idiom varieras som ett talesätt (se figur nedan).



Figur 3

Talesätt eller idiom

Gängse ordböcker i fraseologi har olika åsikter om ordspråken ska räknas in bland de fasta fraserna eller inte (Carter 1987, s. 60 o. Gläser 1986 s. 105). Vi tar med dem, men i första hand sådana ordspråk som är möjliga att variera.

I åtskilliga talesätt ligger ofta något modalt: 'man bör, får, måste, kan' etc. Sådana uttryck förkortas ofta och uppfattas då som idiom. En formalisering som [*man ska/måste/får*] *ta seden dit man kommer* anger alltså att ett ursprungligt talesätt kan varieras något i början eller avkortas och övergå till idiom. I det avkortade uttrycket finns emellertid fortfarande den pekpinne som talesättet explicit uttrycker.

Talesätt av samma typ är:

[*man ska inte*] *döma / skåda hunden efter håren* '[man ska inte] bedöma ngn/ngt enbart på grundval av yttre egenskaper',

[*man ska inte*] *skåda given häst i munnen* '[man ska inte] kritiskt granska en gåva',

[*bli tvungen att*] *ta skeden i vacker hand* '[bli tvungen att] anpassa sig',

[tvingas] bita i det sura äpplet '[tvingas] acceptera något obehagligt'.

Det kan också hända att ett idiom byggs ut till talesätt. Ett idiom som ligger nära talesättet är t.ex. *strö salt i såret / såren [på ngn]* som naturligt kan förlängas: *[man ska inte] strö salt i såret / såren [på ngn]*.

Idiom eller jämförelse

Den som vill använda ett idiom kan få för sig att mottagaren behöver extra påpekanden om att det uttryck som han vill använda inte ska uppfattas bokstavligt utan som en jämförelse. Därför tillgriper många ett *liksom* eller *så att såga* för att underlätta för mottagaren att tolka uttrycket.

I sådana fall har SKO:s redaktion bortsett från utvecklingen; frasen ges avskalad som ett idiom. Ett exempel som "Du förstår att i Sverige går man på *liksom* i ullstrumporna och inser inte att vi är översprungna för länge sen" bör alltså formaliseras som ett idiom: *gå på i ullstrumporna* 'fortsätta som man börjat'.

Även ett vanligt *som* kan fungera som en metaspråklig kommentar, t.ex. i *vara [som] ett rött skynke för ngn* 'vara en källa till stark irritation för ngn' som vi ger i form av ett idiom med variationen inom klammer.

Det kan också hända att ett jämförelseuttryck avkortas till ett idiom. I sådant fall bör uttrycket klassificeras och formaliseras som en jämförelse, eftersom jämförelsen är det primära uttrycket. Jämförelsen *sitta som på nålar* 'vara mycket otålig' kan t.ex. eventuellt ges med *som* inom klammer, vilket skulle visa att uttrycket kan varieras som ett idiom. Att vända en jämförelse till idiom är alltså ett sätt att skapa nya idiom.

Variation inom uttrycket

Den variation som finns inom uttrycket är av tre slag: variation på frasnivå, variation på exempelnivå och kreativ variation. Den förstnämnda bör skrivas in i den formaliserade frasen, den andra kan eventuellt registreras i exempelform. Den sistnämnda däremot ges varken på frasnivå eller exempelnivå, men den kan eventuellt kommenteras.

1. Variation på frasnivå

Många av idiomblockets fraser tillåter en mindre variation inom en viss ram. En sådan variation bör ges i den formaliserade versionen.

a) morfologisk/syntaktisk variation

Någon gång kan numerus varieras. Frasen *strö salt i såret / såren [på ngn]* 'fortsätta att angripa ngn som redan lider' – som i sin betydelse ligger nära ett talesätt – bör sålunda skrivas *[man ska inte] strö salt i såret / såren [på ngn]*.

Ordföljden kan ibland varieras. Ex.: *inte veta på vilket ben man ska stå / vilket ben man ska stå på* 'vara tveksam', även 'vara överlycklig'.

b) lexikal variation

I ett idiomatiskt uttryck kan ofta ett av de ingående orden eller flera bytas ut mot ord av ungefär samma typ. Själva kärnan i frasen är då fast, medan ord i utkanterna av uttrycket är varierbara inom en viss ram. Ofta kan verbet varieras. I frasen *lägga benen på ryggen* 'springa så fort man kan', 'fly', kan verbet *lägga* bytas ut, eventuellt också substantivet *rygg*. Frasen bör skrivas sålunda: *lägga* el. *ta benen på ryggen* el. *nacken*.

Även ett adjektiv kan varieras: *springa [omkring] som yra* el. *skrämda höns* 'vimsa omkring', eller ett substantiv: *stå med bägge fötterna* el. *benen på jorden* 'vara realist'.

c) tillägg/avkortning

Åtskilliga fraser inom idiomblocket har alltså en kärna som inte får ändras, medan orden i utkanterna tillåter en viss variation, om den inte förändrar uttryckets betydelse. Verben är som redan sagts ofta varierbara. De kan t.o.m. strykas när kärnan uppfattas som det egentliga idiomat. Ex.: [*vara*] *femte hjulet [under vagnen]* 'den överflödiga', ofta om den tredje som blir utanför i sällskap med ett par. Här kan också senare delen av uttrycket förkortas. I idiomat *sätta / ha svansen mellan benen* kan t.ex. verben strykas och ett nytt idiom bildas: *med svansen mellan benen* 'skamsen'.

Bland talesätten finns många uttryck som är relativt långa och kan tyckas onödiga att citera i sin helhet: mottagaren förstår ju ändå vad man menar. Ett flerledat ordspråk som *tala är silver, tiga är guld* citeras ofta endast delvis: *tala är silver*. Ordboken bör emellertid ge hela uttrycket men medelst klammer visa på möjligheten till avkortning. Här är man dock fortfarande kvar på talesättens område. Detsamma gäller det rumphuggna *Det är ingen ko på isen* där man saknar fortsättningen *så länge rumpen är i land*; uttrycket betyder snarast 'det är ingen fara', 'man behöver inte skynda sig'.

Ett idiom som *ro fan i land* kan tyckas väldigt speciellt med sin betydelse 'genomföra något som man börjat med fastän det blir svårt' men om man får se början blir det genast klarare: *Har man tagit fan i båten* och fortsättningen *så får man ro honom i land*.

Idiomat *helig ko* 'ngn el. ngt som ej får kritiseras' kan utgöra bas i ett nytt idiom: *slakta en helig ko / heliga kor* i t.ex. "Han är okänslig för det oväsen som uppstår när heliga kor slaktas."

En avkortning kan leda till *avidiomatisering* och åstadkomma rotation i förrådet av fasta fraser. I idiomat *smaka förbjuden frukt* 'njuta ngt som är förbjudet' kan objektet tas ut och behandlas som ett vanligt fast bildligt uttryck: *förbjuden frukt* 'förbjuden njutning'.

2. Variation på exempelnivå

En situationsbunden variation som ändrar uttryckets betydelse kan eventuellt ges på exempelnivå. En synnerligen vanlig situationsbunden variation är utvidgning med adjektivattribut. Så kan t.ex. idiomat *räta ut frågetecknen kring ngt* 'klarlägga ngt' varieras: "Nu rycker frågan fram om Sveriges förhållande till den europeiska gemenskapen. Sedan EG rätat ut sina *svåraste* frågetecken står den nära och ställs konkret."

En annan möjlighet att anpassa ett idiom till situationen är att bilda en sammensätt-

ning. I idiomet *slå i taket* 'inte komma längre' kan *taket* få en förled: "Vi riskerar att upprepa misstagen från 1980-talet då de svenska företagen slog i *kapacitetstaket*." På samma sätt får substantivet *pund* i idiomet *förvalta sitt pund* 'sköta sina tillgångar' en situationsbetingad förled: "Under de gångna krisdagarna har många av oss som är verksamma på morgontidningarna knappt trott våra ögon när vi har sett hur framförallt TV:s nyhetsredaktioner har förvalt sitt *licens- eller reklampund*."

3. Kreativ variation

Många idiom är utgångspunkt för en kreativ skribents utflykter. Ett nytt uttryck skapas, även om man kan ana det gamla i botten. Sådana variationer ska varken anges i grundform eller exempel. Eventuellt kan man på kommentarnivå antyda möjligheterna.

Idiomet *koka soppa på en spik* 'åstadkomma ngt bra av ingenting' inbjuder till variation: "Helge Skoog är inte bara Ture Sventon i de svenska stugorna, han är också den som kokar mustig teater på olika spikar i Stockholms Stadsteaters regi sedan flera säsonger". I exemplet "Särskilt många tappade hakor återfanns på redaktionsgolven den dagen Rapport via en ekvilibristisk övergång ställde riksbankschefens värdering av vår ekonomiska situation mot uppfattningen hos ABAB-vakten i riksbankens huvudentré" känner man igen mönsteridiomet *tappa hakan* 'öppet visa sin förvåning'.

Det finns många möjligheter att variera ett idiom. Därför är det kanske inte alltid så lätt för lexikografen att avgöra vilka variationer som ska tas med i ordboken och vilka som bör anses vara tillfälliga och sorteras bort. Det kan ibland också finnas olika åsikter om vad som är ett uttrycks normalform. Eventuellt kan man undersöka sådant genom enkäter. Det bör emellertid framhållas att det aldrig är förbjudet att variera ett idiom, men användaren bör ha normalformen klar för sig. Och här kan en god ordbok vara till hjälp.

Litteratur

- Anward, Jan & Linell, Per (1976), Om lexikaliserade fraser. I: *Nysvenska studier*, Årg. 55-56, s. 77-119.
- Burger, Harald (1982), *Handbuch der Phraseologie* / von Harald Burger, Annelies Buhofer u. Ambros Sialm, Walter de Gruyter, Berlin.
- Carter, Ronald (1987), *Vocabulary*. Allen & Unwin, London.
- Fleischer, Wolfgang, *Phraseologie der deutschen Gegenwartssprache*. VEB Bibliographisches Institut Leipzig.
- Gläser, Rosemarie (1986), *Phraseologie der englischen Sprache*. VEB Verlag Enzyklopädie Leipzig.
- Moon, Rosamunda (1992), 'There is reason in the roasting of eggs': a consideration of fix-ed expressions in native-speaker dictionaries. I: *EURALEX '92, Proceedings II*, s. 492-532
- Svensén, Bo (1987), *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. Stockholm.

Om datering af ordene i dansk

Anne Duekilde

Det hører med til almindelig nysgerrighed og interesse for sproget at spørge: hvor længe har vi sagt sådan? det føles ikke ligegyldigt om en ordanvendelse er 2 år eller 20 år eller 200 år gammel. For forskningen har oplysninger om ordenes levetid i sproget både sproghistorisk og kulturhistorisk interesse; det kan være nyttigt at vide hvilket sprogligt materiale der var til rådighed på et bestemt tidspunkt, og oplysningerne om en bestemt epokes nye ord siger noget om kulturindflydelse og om epokens sproglige aktivitet.

Ved visse ordbogstyper er datering det centrale emne. Det gælder ordbøger over nye ord i en vis periode og de sjældnere ordbøger over ord der forældes i en vis periode. Dateringsforskningen kan resultere i en kronologisk ordbog, som viser forandringerne i ordforrådet gennem en vis periode¹. Dateringsspørgsmålet indgår også i de store nationalordbøger, og ud fra deres materiale har forskellige mindre ordbøger bragt dateringsoplysninger.

Da min baggrund for at vove mig ind på dette omfattende emne er arbejdet som redaktør ved Supplement til Ordbog over det danske Sprog (ODS-S.Iff. 1992ff.), vil jeg specielt se på hvordan (begyndelsen af) ODS og ODS-S bidrager til dateringen af ordene; desuden vil jeg berøre forholdet mellem etymologisk og semantisk datering samt visse problemer ved førstebelæg og sidstebelæg.

ODS som dateringskilde

ODS er planlagt som en stor historisk citatordbog der skulle blive hurtigt færdig, nemlig på cirka 20 år. ODS ville i 15-17 bind beskrive samtids sproget og desuden vise hvordan dette havde udviklet sig siden 1700. Ambitionsniveauet blev lagt betydeligt lavere end i Svenska Akademiens Ordbok, som var nået frem til bogstav D i de år begyndelsen af ODS blev redigeret. Første halvbind af ODS udkom 1918, og halvdelen af alfabetet blev faktisk redigeret på cirka 10 år. Sidste egentlige ordbogsbind udkom i 1954 og det afsluttende forkortelsesbind i 1956; det blev til i alt 28 bind. Den ældre nydanske tid, 1500-1700, behandles ikke i ODS; denne periode havde netop fået en leksikografisk behandling i Otto Kalkar, Ordbog til det ældre danske sprog (1300-1700), hvis store supplementsbind, værkets bind V, blev afsluttet i 1918.

ODS er både en samtidsordbog og en ordbog over de to foregående århundreder. Fra hele perioden blev der gennemgået tekster af mange slags, men interessen var ikke lige-

ligt fordelt mellem fortid og nutid. Nutidssproget var hovedsagen, og derfor blev der oprindelig ikke lagt særlig vægt på ordenes datering. Det var dog en fast regel at det ældste kendte belæg på et ord eller en ordanvendelse altid blev medtaget. De skønlitterære hovedforfattere lige fra Ludvig Holberg til Andersen Nexø blev godt repræsenteret i seddelsamlingen, men fra den ældre tid blev forholdsvis få andre tekster gennemgået, hvorimod samtidens bedste opslagsværker inden for alle fagområder blev udnyttet. I den senere del af ODS blev der inddraget flere ældre tekster.

Dateringsoplysninger kan i ODS fremtræde enten som en redaktionel oplysning eller i form af kildemateriale. Man kan skelne mellem 3 slags redaktionelle oplysninger:

1. Oplysninger om forekomst i en sprogperiode før 1700.
2. En redaktionel dateringsoplysning ved ord eller ordanvendelser der først er kendt i ODS' periode, fx: *I. Ballade*, digt: "laant i slutn. af 18. aarh. fra ty. ballade", *II. Ballade*, løjer: "optaget i kbh. i slutn. af 1860'erne fra parisisk ballade, spadseretur uden bestemt maal, kun for at drive ell. more sig", *Barndoms-*: "først alm. i 19. aarh., jf. at VSO.(1793)² ingen eksempler har", *benytte*: "dannet i 1. halvdel af 18. aarh. efter ty. benutzen, benützen, af ældre nytte".
3. Et vinkeltegn "<" foran forkortelsen for en ordbog er en indirekte dateringsoplysning. Det angiver at ordet (eller ordanvendelsen) ikke er medtaget i vedkommende ordbog; det er endda forklaret med ordene: yngre end, i tillid til at ordbøger altid afspejler den eksisterende sprogbrug. Fx: *afklare* 2, bringe klarhed m.h.t. noget: < VSO.1793, *Alvorsmand*: < MO.1859 (førstebelæg 1881, ODS-S 1818), *aflaste*: < S&B.1883, *berolige*: < vAph.(1764).

Hovedreglen for det nye ordstof fra ODS' periode 1700ff. er at dateringen må aflæses af kildematerialet. Det samme gælder for det ordstof som er gået af brug i løbet af perioden; yngste belæg må findes i kildematerialet. Det er derfor en alvorlig mangel ved ODS at der ikke er sat årstal på ethvert belæg; og denne mangel kan et supplement ikke afhjælpe. Læseren får altså normalt ikke mulighed for at aflæse ordets alder eller levetid direkte, men skal i reglen gå omvejen over forkortelsesbindet. Hvis citatet er fra en udgave hvis trykår er meget forskelligt fra tekstens år, fx samlede eller udvalgte værker, kan læseren endda stå uden anden datering end forfatterens fødsels- og dødsår.

ODS er en betydningsordbog. Kildematerialet er derfor grupperet betydningsmæssigt, og under hver gruppe er det ordnet kronologisk. Man skal se sig godt for, hvis man vil finde det ældste belæg på ordet som helhed, for det står ikke altid tidligt i artiklen, og det kan stå i en underordnet citatgruppe.

Trods de nævnte forhold er ODS vor vigtigste kilde for dateringsoplysninger inden for den periode ordbogen skal dække.

ODS-S som dateringskilde

ODS-S har til formål at supplere ODS' oplysninger om det danske sprog i perioden 1700-1955. Modsat ODS, og de store udenlandske ordbøger som var forbilleder for

ODS, har ODS-S altså sproget i en afsluttet periode som sin genstand, og ikke det dagsaktuelle sprog. Holdningen til det 20. århundrede bliver derfor tilbageskuende, og et "nu" indgår ikke i redaktionssproget. Så vidt det er muligt bliver de supplerende oplysninger fremstillet i de redaktionelle rammer som er givet i ODS. Dette gælder både for de artikler der supplerer en ODS-artikel og for de helt nye artikler. I ODS-S I er der 10.356 opslagsord, og cirka halvdelen supplerer en ODS-artikel. I lidt over 500 af disse står der et "+ ældre: "; på denne måde er et sprog- eller kulturhistorisk markant ældre belæg gjort synligt.

En gennemgang af disse tilbagedateringer inden for ODS' periode viser en klar dominans af tilfælde hvor ODS har ældste belæg omkring 1900. Her i begyndelsen af alfabetet, ODS-S I går fra A (inklusive Aa) til Bh, er ODS i allerhøjeste grad samtidsordbogen - som ganske vist var historisk, men som ikke var i besiddelse af tilstrækkeligt med informationer om fortiden. Ældste belæg er påfaldende ofte Salmonsens Konversationsleksikon, 1. udg. 1893-1911; eller det er fra tidens bedste faglitteratur inden for landbrug, teknik, handel, medicin osv. Selv i tilfælde hvor ældre opslagsværker eller ældre udgaver af benyttede værker kunne have været efterset, har man oftest ladet sig nøje med samtidens standardværker. Til en vis grad findes denne svaghed gennem hovedparten af alfabetet, specielt er fremmedordene i alt for høj grad forbigået ved excerperingen til ODS.

Tilbagedateringer fra ca. 1900 til det anførte år: *Aarebrok* (> 1810), *Aarring* (i en træstamme, > 1858), *Afgangsprøve* (> 1856), *Akvarium* (> 1858), *Allitteration* (belæg fra sprogforskeren Otto Jespersen, > 1816), *ambulatorisk* (> 1850), *Anlægs kapital* (> 1851), *anomal* (> 1838), *Association* (filos., psyk. > 1824), *Bakterie* (> 1868), *Barometerstand* (> 1826), *Begonia* (> 1834), *Benkul* (> 1861), *Beskyttelsestold* (> 1840), *Bevægenerve* (> 1840). Nogle eksempler fra det mere almene tale- og skriftsprog: *Abort* (billedligt om et mislykket åndsprodukt, > 1845), *akademisk* (i betydningen formel, regelbunden, > 1735), *Akkordarbejder* (> 1816), *Akrobat* (> 1828; ODS oplyser, vel for at bøde på sit utilfredsstillende kildemateriale: "først brugt i 19. aarh."), *altomfattende* (> 1811), *amerikanisere* (> 1833: et amerikaniseret Europa), *analytisk* (> 1775), *Anneksion* (om en stats indlemmelse af et landområde, > 1864) *annekttere* (om en stat, > 1870), *annoncere 2* (bekendtgøre gennem annoncer, > 1825), *appetitvækkende* (> Søren Kierkegaard, 1840'erne), *Arbejderurolighed* (> 1848), *arbejdsbesparende* (> 1857; jf. *arbejdssparende* 1859 > 1827), *Arrigtrold* (> 1810), *Badulje* (støj, spektakel, > 1772), *Balfaldera* (> 1858), *Bandbulle* (> 1810), *barbarisere* (> 1843), *Batteri 3* (spøg. om en samling flasker, > 1864), *Bavian* (som skældsord, > 1787), *bedugget* (beruset, > 1779), *begrave 3.2 slutn.* (m.h.t. lovforslag olgn., > 1858), *Ben 2.3* (om bierhverv, > 1858).

Tilbagedateringer fra ODS-belæg ældre end ordbogens samtid: *afgaa 2.1*, få ell. tage sin afsked, VSO.1793 > HolbO., *afskedige 2*, opsige fra et arbejde, (< vAph.(1764)) VSO.1793 > HolbO., *Allongeparyk* 1869 > 1740, *amusant* 1800 > 1751, *analysere* ca. 1780 > HolbO., *Ballastsand* 1860 > 1788, *Balsam*, i bet. lindring, trøst, ca. 1830 > 1732, *III.bar 3.2*, fx som bare pokker, ca. 1840 > 1785, *Barberkost* MO.1859 > 1797, *give barnet navn* (sp.1160,2), sige noget rent ud, ca. 1800 > 1701, *barrikadere* JBaden.FrO.1820

> 1765, *Bassin* 1808 > 1765, *II.basta*, interj., 1807 > 1749, *befuldmægtige* VSO.1793 > HolbO., *Benrad*, om døden, 1824 > 1776, *Berøringspunkt* 1843 (billedl.) > 1799 (egl.); jf. ordet *Berøringsflade* S, 1805, den billedl. brug belagt ca. 1935, *betalbar* 1873 > 1818, *Beton* (< MO.1859) 1873 > 1868, i sammensætningen *Betongrus* 1838.

En tilbagedatering med ret få år har især interesse i nærheden af ordets eller ordanvendelsens formodede fremkomst. Nogle eksempler: *alkoholisere*, mht person, 1908 > 1889, *Alkoholist* 1917 > en fremmedordbog 1899; ordet er for øvrigt trængt af *Alkoholiker* S, 1905, *Arkæologi* 1815 > 1807, *Automatkafé* 1912 > 1902, *Bajer*, bajersk øl, 1877 > 1857; ordet har udkonkurreret substantiveringen *bajersk*, (fx "sit Kruus Bajersk") som er brugt i 1850'erne, *Ballon*, i bet. haløj, ballade, 1870'erne > 1859, *Banjo* 1894 > 1883, *Barselheber* ca. 1850 > 1827, *Bersærkergang* beg. af 1800-tallet > ca. 1780.

Det modsatte af tilbagedatering kan vel kaldes fremdatering, det vil sige en angivelse af yngre (senere) belæg end hidtil kendt, og en sådan fremdatering må ifølge sagens natur angå ordanvendelser som ikke mere er i brug, ellers var der jo ikke tale om belæg yngre end hidtil kendt.

Ordanvendelser der i ODS er betegnet med kors har som sidste belæg et citat fra det 18. eller 19. århundrede. Hvis der nu kan suppleres med et belæg der er betydelig yngre, er ordet fremdateret: *Beliggende*, beliggenhed: sidstebelæg H.C. Andersen > 1895, *Be-slutning 1*, ende, afslutning: sidstebelæg før midten af 19. årh. > 1886, *bevant 1*, øvet, erfaren: sidstebelæg 1816 > 1932 (en bevant murersvend; i en stillingsannonce, som eventuelt er dialektpræget).

Et yngre citat i ODS-S er ikke altid en fremdatering, ofte er det blot et yngre eksempel på noget der stadig er fuldt brugeligt. Det er medtaget fordi ordbogsværkets nutid bør være repræsenteret i kildematerialet (men ikke nødvendigvis slutningen af ODS-perioden). Eksempler på citatmangel af denne art i ODS: *alenlang*, overf.: yngste belæg i ODS er fra midten af 19. årh., + yngre: 1955, *almindeligvis*: H.C. Andersen yngst, + yngre: 1953, *Angstskrig*: kun ordbøgerne VSO.1793 og MO.1859, + yngre: 1954.

Ved de egentlige fremdateringer er det nødvendigt med en brugsrestriktion. Det kan i ODS-S hedde: "+ yngre (men aftagende brug i 20. årh.)" eller: "+ yngre (men ikke alm. i 20. årh.)".

En vis del af det ordstof der var almindelig brugt da begyndelsen af ODS blev redigeret, er i slutningen af ODS-perioden godt på vej ud af sproget. Dette forhold kan give anledning til bemærkningen "aftagende brug i 20. årh."; et påfaldende sent citat medtages gerne, for at vise ordenes sejlivedhed. Eksempler på ord med aftagende brug i 20. årh. er: *agtbar* (som får tilføjet et Karen Blixen-citat), *Arbejdskone*, *-kvinde*, *Automobilist* (kan ses i Rigsdagstidende 1946), *basune* (hvorimod *basunere*, der ifølge ODS var mindre brugt, det var måske ret nyt omkring 1900, får bemærkningen "øget brug i 20. årh.").

Nu kan man spørge om dateringen i ODS-S, også dateringen af de ord som er nye i forhold til ODS, er til at stole på. - Det er den naturligvis ikke (jævnfør de ndf. nævnte pro-

blemer vedrørende førstebelæg og sidstebelæg). Mange af ordene må kunne findes 10-20 år tidligere, men næppe mange af de daterbare ord er 50-100 år ældre end vort ældste belæg (når vi ser bort fra tiden før 1700, som vi ikke ved nok om og som ligger uden for ordbogens periode). Og med hensyn til sidstebelæg kan mange af de ord der er angivet som kun tilhørende ældre sprog, dvs sproget før ca. 1900, utvivlsomt påtræffes senere end vi har kunnet vise.

Datering af etymologiske eller semantiske enheder

Spørgsmålet om hvornår ordene dukker op i et bestemt sprog kan siges at høre med til de etymologiske oplysninger. Elof Hellquist medtager i sin Svensk etymologisk ordbok, hvis 1. udgave er fra 1922, årstal for ordenes tidligste optræden i svensk, så vidt det har været muligt. I Niels Åge Nielsen, Dansk Etymologisk Ordbog, 1966 (og senere udgaver) er ordets alder i dansk ikke anført direkte, men der er givet oplysning om forekomster i runedansk, gammeldansk eller ældre nydansk.

Hvis man i en etymologisk ordbog vil se hvornår de ord vi bruger i dag er kommet ind i sproget, så bliver svaret måske misvisende. For det man ønsker at vide noget om er de kendte ordanvendelser, altså ordet i bestemte semantiske anvendelser, snarere end det er den etymologiske ordform. En sådan semantisk ordanvendelse kan være en ordbetydning, en enhed af flere ord (idiom, term) eller en produktiv orddel (affiks, sammensætningsled).

De følgende eksempler skal vise at en rent etymologisk datering ofte rammer ved siden af det moderne ords hovedbetydning (jeg følger i eksemplerne ODS-artiklernes struktur og supplerer med ODS-S-artiklernes oplysninger):

Ambulance 1 feltlazaret ell. -sygevojn. 1837ff. +2 civil sygevojn. 1861ff. +3 billedl. 1932ff.

annoncere 1 bekendtgøre. 1700ff. 2 bekendtgøre gennem annoncer. 1825ff.

annektere 1 tilknytte. ca.1700ff. 2 om stat: sætte sig i besiddelse af et landområde. 1870ff.

anløbe 1 bevæge sig mod et mål. 1.1 nærme sig i løb ell. med fart. VSO.(1793), ikke brugt o. 1900. 1.2 sejle ind til (en havn olgn.) for et kortere ophold. 1807ff. +|| (fagl.) om andet end skibe. 1934ff. +1.3 anfalde; angribe. ældre nydansk og 1838. 2 om visse overfladeforandringer (ved metaller). 2.1 (især tekn.) intr.: miste den oprindelige farve og glans. 1777ff. 2.2 (tekn.) trans.: fremkalde en farveforandring på (et metals overflade). 1802ff. 2.3 part. *anløben* brugt som adj. om metaller: 1700ff. || overtrukket med skimmel olgn. om vin: Moth.(1700), ikke brugt o. 1900. || (især naturv.) svagt farvet. 1902ff. || (ofte spøg.) om virkninger af alkoholnydelse. om næsens farve: slutn. af 19. årh. let beruset: 1880ff. || ikke ren ell. frisk; angrebet; let fordærvet. fagl., fx om træ: 1900ff. om noget i moralsk henseende let fordærvet: 1880ff.

Det etymologiske ord *anløbe* er altså kendt i ældre nydansk i betydningen 'angribe' (i Dansk Etymologisk Ordbog er der ikke betydningsangivelse). I almensproget i dag bruges ordet især om skibe der sejler ind til en havn; participiet *anløben* eller *anløbet* har skilt

sig ud og opfattes nu som et selvstændigt ord. Eksemplet skal illustrere at man ikke kan "løbe an på", at den etymologiske ordform i sin ældste anvendelse ligner de moderne betydninger.

Eksempler på enheder af flere ord: Ordforbindelsen (idiomet) *løbe an på*, med betydningen 'regne med at noget vil ske eller forholder sig på en vis måde', er måske samme etymologiske enhed som anløbe, men jo en selvstændig semantisk enhed; under II.løbe 22.3 i ODS er det ældste belæg fra ca. 1880. Idiomerne *i lange baner* og *i brede baner*, begge med betydningen 'i mængde', var ifølge ODS almindelige i daglig tale o. 1900; det sidste udtryk har ikke holdt sig i sproget. Faglige termer som *betinget dom*, 1911ff., og *+ betinget refleks*, 1932ff., bør naturligvis dateres som selvstændige enheder.

Produktive orddele med en særbetydning er fx: *-apostel* i ord som afholdsapostel, sandhedsapostel, sundhedsapostel, der er kendt fra 1911 (afholdsapostel); *-bank* om et depot af andet end penge er kendt i dansk siden 1943, ældst i blodbank. Også disse ord-dannelsesmuligheder er det rimeligt at få dateret.

Problemer vedrørende førstebelæg og sidstebelæg

En sikker datering forudsætter naturligvis en omfattende registrering af ordforrådet gennem tiderne. En forbedret registrering af dansk vil indtil videre medføre tilbagedateringer og fremdateringer. Men der er en del særlige forhold at betænke ved efterforskningen af ordenes liv og død.

En vurdering af ordets art:

Visse sider af ordforrådet lader sig vanskeligt datere, eller sagt på en anden måde: det foreliggende førstebelæg vil næppe ligge tæt på ordanvendelsens oprindelse. Dette gælder

- a. De ældste ord i sproget, arveordene, som går længere tilbage end kilderne rækker.
- b. De udprægede talesprogsord fra både ældre og nyere tid. En stor del af dialekternes ordforråd er kun kendt fra optegnelser inden for de sidste 100 år, men er naturligvis ældre. Slangudtryk kan leve kort eller de kan leve længe, det er svært at vide, og kil-dematerialet svigter.
- c. De selvfølgelige eller tilfældige orddannelser som står til rådighed i sproget. Et ord som *Beslutningskraft* anså ODS-redaktøren for dødt, modsat beslutningsdygtighed og -evne; ODS-S viser et 1950-citat, og ordet kan sandsynligvis også træffes i begyndelsen af århundredet; det er ikke almindelig brugt, men kan nydannes af den der vil.
- d. Visse låneord. Hellquist gør i sin indledning opmærksom på at han har undladt at datere en del romanske eller klassiske låneord, da deres ældste optræden ofte er rent tilfældig, som citatord, og måske ikke har direkte historisk sammenhæng med den senere, faste brug.

Den anden yderlighed mht dateringsmuligheder er de kulturord som kan dateres helt præcist. Et kendt eksempel på et ord med fødselsattest er ordet *bil*, der blev foreslået til

en dagbladskonkurrence i 1901. Inden for videnskab og teknik er der en række internationale ord som er dannet og defineret på kendte tidspunkter af bestemte ophavsmænd; også visse officielle ord vedrørende samfundsforhold er dannet eller indlånt på bestemte, kendte tidspunkter.

En vurdering af dateringskilderne:

Det bedste kildemateriale er naturligvis citater fra løbende tekst hvor tid, sted, situation og sprogbruger er kendte størrelser. Ordbøger kan være farlige som dateringskilde, da ordbogsord ikke altid afspejler den faktiske sprogbrug. Oversættelsesordbøger har ofte konstrueret ord ud fra det fremmede ordstof³. Fremmedordbøger, især de ældre, støtter sig på udenlandske ordbøger og har ikke altid været så nøjeregnende med om ordene var at finde i danske tekster.

Øjenvidneberetninger er nyttige. ODS citerer ofte de ældre ordbøger for bemærkninger som "nyt ord" (se fx u. betone) eller bemærkninger om at et ord er ved at udgå. ODS selv er et øjenvidne til sproget i redaktionstiden. I 1955 fik vi Dansk Sprognævn, som bl.a. har til opgave at holde øje med sprogforandringerne. På grundlag af sprognævnets iagttagelser har vi fået ordbogen Nye ord i dansk 1955-75 (1984), redigeret af Pia Riber Petersen (nu: Pia Jarvad). En registrering af de ord som mange vil anse for døde, men som alligevel jævnlige dukker op, har derimod ikke interesseret sprognævnet ret meget.

En vurdering af frekvens og stilværdi:

Begrebet datering rejser spørgsmålet om kriterier for hvornår en ordanvendelse kan siges at tilhøre sproget. Nye ord i dansk 1955-75 indeholder ifølge indledningen (s. 11) "nye ord og vendinger som er blevet almindelige i perioden 1955-75"; det er altså ikke nødvendigvis absolutte førstedateringer. Hertil svarer forekomsten af oplysninger som: "nu i almensproget" (u. socialgruppe), "øget brug" (u. ulækker), "øget brug i almensproget" (u. risikogruppe, snalret). Ord der fx kun er kendt inden for dialekterne tilhører naturligvis dansk, men dog ikke det fællesdanske sprog. Det samme gælder ord fra det interne fagsprog.

Ord kan komme ind i sproget pludseligt, og ord kan siges at komme langsomt; denne skelnen forudsætter at man tager hensyn til frekvens og stilværdi. Nogle ord kan leve et kort og hektisk liv som modeord.

Ord kan gennem en meget lang periode være på vej ud af sproget; måske bliver de stående i en niche og fungerer som højtidelige eller som uhøjtidelige, spøgende ord, eller de giver teksten et gammeldags præg.

Dateringsforskning er et område hvor der endnu ikke er sket ret meget. En sådan forskning kan have interesse for både sprog-, kultur- og litteraturhistorien.

Noter

1. Robert and Clarence Barnhart: *The Dictionary of Neologisms*, i: Wörterbücher, Dictionaries, Dictionnaires. Red. af F.J. Hausmann et al. Berlin & New York. II. 1990. s. 1159-66. - Oscar Reichmann: *Wörterbücher archaischer und untergegangener Wörter*. smst. s. 1153-58. - Wilhelm Kesselring: *Das chronologische Wörterbuch*. smst. s. 1342-48.
2. Følgende kildeforkortelser forekommer i eksemplerne: *HolbO.* = Aage Hansen og Sv. Eegholm-Pedersen: *Holberg-Ordbog*. I-V. 1981-88. - *JBaden.FrO.* = Jacob Baden: *Ordbog over de af fremmede Sprog laante Ord*. 2. udg. 1820. - *MO.* = C. Molbech: *Dansk Ordbog*. 2. udg. I-II. 1859. - *S* = der henvises til en artikel i ODS-S. - *S&B.* = Th. Sundbye & E. Bartiel: *Dansk-norsk - fransk Haand-Ordbog*. I-II. 1883-84. - *vAph.* (1764). = H. von Aphelen: *Kgl. Dansk Ord-Bog*. I. Dansk og Tydsk. 1764. - *VSO.* = *Dansk Ordbog* udg. under Videnskabernes Selskabs Bestyrelse. I-VIII. 1793-1905.
3. Bo Ralph: *The older dictionaries as sources for Nordic language history*, i: J. Louis-Jensen & J.H.W. Poulsen (eds.): *The Nordic Languages and Modern Linguistics* 7. Tórshavn 1992 s. 493-509.

Genbrug af Ordbogsdata: En dansk Ordbank

Dorthe Duncker og Ole Norling-Christensen

Den Danske Ordbog (DDO) i 6 bind skal udkomme 1998-99. Den skal beskrive sproget efter ca. 1950 med hovedvægt på perioden 1983-92, ordbogens såkaldte *primærperiode*.

Datamatiske hjælpemidler benyttes i stort omfang som støtte for det redaktionelle arbejde, og vi har samlet en væsentlig del af ordbogens kilder i to store databaser. Den ene er DDO's tekstkorpus på 40 mio. ord, der dækker primærperioden (Kristensen 1994). Den anden er *Ordbanken*, der omfatter bearbejdede data fra en række eksisterende ordbøger, hvis oplysninger er af betydning for DDO. Det er Ordbanken, og vekselvirkningen mellem denne og korpuset, som er emnet for denne artikel.

Tanken om at samle og koordinere de leksikalske oplysninger om dansk, som allerede foreligger, men er spredt på mange forskellige kilder, er ikke ny. Den var udgangspunkt for det forskningssamarbejde, som indledtes i 1982 og siden fik navnet DANLEXgruppen (DANLEX 1987a); den vandt tilslutning såvel i den danske ordbogsverden (DANLEX 1987b), som i bredere kredse af professionelle ordbrugere (DANLEX 1989); og den var en selvfølgelig og nødvendig del af den plan (Dansk Ordbog 1989), som blev grundlaget for DDO-projektet med dets redaktionsperiode på kun 6 1/2 år.

Mens sproget i primærperioden i det væsentlige skal beskrives på grundlag af korpuset, er hovedkilderne til tiden før 1983 Dansk-fransk Ordbog (B&H) og Dansk-engelsk Ordbog (V&B), hvis førsteudgaver begge byggede på Ordbog over det Danske Sprog 1-28 (ODS). Sammen med Retskrivningsordbogen (RO), den officielle danske norm for stavning og bøjning, samt Dansk Sprognævns Ordregister (DSO) (nøglen til en samling på henvend en million excerpter, som Sprognævnet har indsamlet siden sin oprettelse i 1955), har disse ordbøger gennemgået en omfattende datamatisk behandling, der har resulteret i en database over mere end 300.000 danske ord med angivelse af ordklasse, ortografiske varianter, bøjnings- og orddannelsesoplysninger, betydningsdifferentiering og eksempler på mulig kontekst.

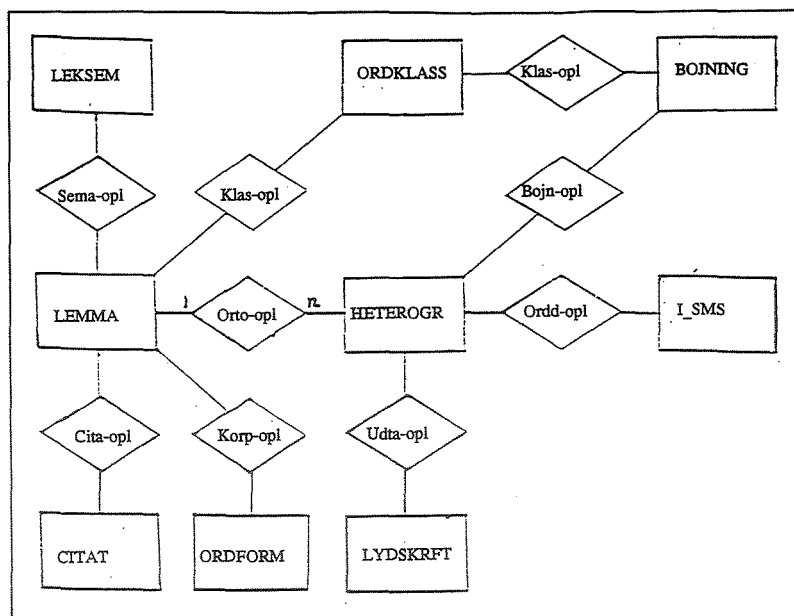
Udover funktionen som kilde til ord, betydning og brug før 1983, og til korrekt dansk stavning og bøjning, har Ordbanken endnu et væsentligt formål, nemlig at identificere samtlige de ordformer, som forekommer i korpus, dvs. henvføre dem til en eller flere bøjningsformer af et eller flere lemmer. Dette er nemlig en (nødvendig, men ikke tilstrækkelig) forudsætning for at kunne søge præcist i korpus efter alle former af et givet lemma. Desuden giver det et vigtigt bidrag til den ordbogsmæssige beskrivelse af, hvordan den faktiske sprogbrug forholder sig til normen.

Materialet fra ordbøgerne bliver derfor suppleret med de "nye ord", som findes i korpus, og med de yderligere stave- og bøjningsoplysninger som kan udledes af korpus. Si-

den hen vil også udtaleoplysninger blive tilføjet, samt formentlig oplysninger om (hel eller delvis) synonymi.

Ordbankens struktur og indhold

Ordbankens opbygning fremgår af entitet-relationsdiagrammet i fig. 1 (jf. fx DANLEX 1987a: 234-239). Hvert rektangel repræsenterer en entitet og hver rombe en relation mellem to entiteter. Det anvendte databasesystem er Paradox (version 4.0; Borland International Inc.). Med få undtagelser er hver entitet og hver relation realiseret som en særskilt tabel, hvortil der er knyttet et eller flere indekser. Den typiske entitetstabel rummer kun to felter, nemlig en entydig identifikation af posten i form af et nummer og et indholdsfelt. En relation sammenknytter to bestemte entiteter, eller evt. én entitet med sig selv. En post i en relation rummer derfor først og fremmest felter med reference til nummeret på en bestemt post i hver af de entiteter, som relationen knytter sammen. Hertil kan komme et eller flere attributter, felter som nærmere bestemmer arten af den pågældende relation. Ordbankens relationer indeholder foreløbig ét attribut, nemlig feltet *Kilde*, som redegør for, hvilken ordbog eller anden kilde den oplysning, relationen udtrykker, stammer fra. Denne måde at strukturere databasen på (den er, i databaseterminologi, fuldt normaliseret) gør det let at tilføje nye relationer mellem ordene, fx synonymi, etymologisk slægtskab, eller referencer mellem sammensatte ord og de led, de er sammensat af.



Figur 1. Ordbankens opbygning af entiteter (rektanglerne) og relationer (romberne). De særlige navne skyldes reglerne for fil-navngivning: højst 8 tegn; ingen danske særbogstaver

Entiteterne

Centralt står LEMMA, den enhed ordbogsartiklen skal beskrive. Det er en abstrakt enhed, som ortografisk i en del tilfælde kan realiseres ved forskellige (grund)former, her kaldt HETEROGRAfer. Til LEMMAet er desuden knyttet CITATer, LEKSEMer, ORDFORMer og en ORDKLASSe.

Til hver HETEROGRAf er knyttet: en BOJNING (oplysning om bøjningsmønster, hvorudfra alle former kan genereres), en eller flere orddannelsesoplysninger (I_SMS), der angiver i hvilken form 'heterografen' indgår som førsteled i sammensætninger; hertil vil komme en eller flere UDTALer, dvs. fonetisk transskription af grundformen.

CITATerne er de danske udtryk fra V&B og B&H.

LEKSEMerne er først og fremmest V&B/B&H's betydningsadskillende oplysninger; hvis der er anført flere betydninger, medtages en eventuel nulmarkeret (første) betydning også.

FORMerne er fortegnelsen over samtlige forekommende ordformer i korpus med reference til de LEMMAer, som formerne iht. rent morfologiske kriterier kan henføres til.

ORDKLASS er blot en nummereret liste over ordklassernes navne; relationen Klas-opl, som knytter den til LEMMA og BOJNING er ingen tabel, men blot udtryk for, at ordklassen i disse to tabeller angives ved dens nummer i ORDKLASS-tabellen.

De fleste af relationerne bliver *mange til mange relationer* (n-m); for overskuelighedens skyld fastholdes dog lemma-heterograf relationen som en 1-n relation; det indebærer fx, at selv om "spise vb." og "spise sb." har samme grundform, har de hver sin post i heterograf-tabellen.

Fra ordbøger til Ordbank

Ordbøgerne, som udgør Ordbankens hovedkilder, foreligger alle i maskinlæsbar form, men er ikke som udgangspunkt sammenlignelige. Hver kilde har derfor gennemgået en forbehandling, hvorunder samtlige relevante oplysningstyper er blevet bestemt og omsat til et fælles format. Vi har i denne forbindelse delt artikler, som indeholdt mere end ét LEMMA, fx "svejs^e, -ede, i teknisk sprog også -te." (RO) og "øl -let, øl -lene; om flaske e.l. med øl: øllen, øl el. øller, bf. flt. ølle(r)ne." (RO).

Den første og centrale kilde i opbygningen af Ordbanken er RO. Derefter følger DSO, B&H og V&B. RO fungerer som bankens ortografiske lovtekst. Det vil sige, at hvis vi finder det samme opslagsord med samme oplysninger om ordklasse og køn, men med forskellige bøjningsoplysninger i en senere kilde, så sejrer foreløbig RO's bøjningsoplysninger. Fx angiver RO, B&H og V&B hver sin bøjningsoplysning ved *krysantemum* ("-(m)en, -(m)er el. krysantemum el. krysantemer, bf. flt. krysantemum(m)e(r)ne el. krysantemerne." (RO); "(c) (- el. -mer)" (B&H); "en -mum(m)er el. pl. d.s. el. -mer" (V&B)), men kun oplysningerne fra RO bliver indført i banken.

RO indeholder to typer opslagsord: dem med bøjningsoplysninger og dem med lodret

streg. Den lodrette streg markerer, at bøjningsoplysninger kan søges under sidsteleddet som opslagsord. Disse to typer kan imidlertid kombineres, således at et opslagsord både har lodret streg og bøjningsoplysninger. Om denne type siger RO: "Oplysninger om bøjning mv. er kun anført, hvis de afviger fra oplysningerne under sidsteleddet som opslagsord, eller hvis der er mere end ét opslagsord med samme stavemåde som sidsteleddet" (RO: 16). Om det er den ene eller anden årsag, der har motiveret oplysningerne, kan man erfare ved at slå op under sidsteleddet. Ved fx *kobber/stik*, der har anført bøjningen "-ket, kobberstik -kene", kan man slå op under *stik* og finde to homogرافnumre: "1. *stik* -ket, stik -kene" og "2. *stik* ubøjel. to.", mens man ved fx *kobber/stikning* med bøjningsoplysningen "-en" kan slå op under *stikning* og se, at oplysningen hér afviger fra sammensætningens: "-en, -er".

Opslagsord uden lodret streg lagres hele i Ordbanken med deres oplysninger. Op-slagsord med lodret streg, men uden bøjningsoplysninger, lagres ved at oplysningerne under sidsteleddet som opslagsord kopieres; fx *kogelflæsk*, *kogelgrej* og *kogeljomfru*, der forsynes med bøjningsoplysningerne fra hhv. *flæsk*, *grej* og *jomfru* som opslagsord. Op-slagsord, som både har lodret streg og bøjningsoplysninger, lagres med de anførte op-lysningsord. I de tilfælde, hvor sidsteleddet ikke er en homogراف, registreres det desuden, at det aktuelle sidsteled kan ændre bøjning i forbindelse med sammensætning. Fx angiver RO ved *hår* som opslagsord bøjningen "-et, hår -ene". Blandt sammensætningerne med *-hår* som sidsteled findes *fimrelhår*, *knurlhår* og *øjenhår*, hvis bøjning kan søges under *-hår* som opslagsord. Man møder imidlertid også sammensætninger med *-hår*, hvor bøjningen afviger fra *hår* som opslagsord, fx *afrolhår*, *pagelhår* og *plyslhår*, som har bøjningsoplysningen "-et". Denne type oplysning - at *-hår* i sammensætninger med betydningen "frisur" mister flertal - kan anvendes, når bøjningen for ord fra de følgende kilder skal bestemmes.

Efter RO er den næste kilde, som behandles, DSO. Strukturelt kan ordene fra DSO sammenlignes med RO's opslagsord med lodret streg, og de lagres i Ordbanken efter samme principper som disse. Hvis sidsteleddet er en homogراف, kan bestemmelsen ikke foregå fuldautomatisk, men må støttes ved manuel stillingtagen i det enkelte tilfælde. Når dette arbejde er tilendebragt, og både RO og DSO er færdigbehandlet, indeholder Ordbanken godt 230.000 forskellige lemmaer.

Den tredje kilde, som inddrages, er den danske side af B&H. B&H behandles før V&B, fordi B&H anfører ordklasse og køn ved sammensætninger, hvilket V&B ikke gør konsekvent. Disse oplysninger er væsentlige, da de kan spare en del håndarbejde i forbindelse med entydiggørelse af homografe sidsteled. Fx vil *hovednøgle* på denne måde kunne behandles fuldautomatisk, idet B&H anfører ordet med sammensætningsgrænse *hovednøgle* og køn, således at det er muligt fuldautomatisk at vælge mellem de to homogرافnumre i RO "1. *nøgle* -n, -r (til lås)" og "2. *nøgle* -t, -r (garnnøgle)" og tilknytte oplysningen fra "1. *nøgle*" til *hovednøgle*. I V&B findes *hovednøgle* med sammensætningsgrænse, men uden oplysning om ordklasse og køn. Da ordet allerede er bestemt i den forrige kilde og ikke er en homogراف, kan det uden videre genkendes automatisk.

Umiddelbart kan det forekomme overdrevet, at vi gør os så store anstrengelser for at genkende et ord, vi allerede har lagret i banken. Det skyldes to forhold: For det første er det nødvendigt at kunne bestemme de enkelte opslagsord i de forskellige ordbogskilder entydigt, således at vi ikke kommer til at lagre det samme ord flere gange, bare fordi vi ikke har været i stand til at tolke kildernes bøjningsoplysninger korrekt. For det andet ønsker vi også at genbruge og lagre artiklernes betydningsspecifikationer (LEKSEMerne) og danske vendinger og udtryk (CITATerne) fra B&H og V&B. Disse oplysninger skal bruges under opbygningen af artikelskeletter, der i første omgang skal fungere som støtte for redaktionsprocessen. Det er derfor vigtigt, at vi med sikkerhed kan knytte disse oplysninger til det rette LEMMA i Ordbanken.

Opslagsordene fra B&H og V&B lagres i øvrigt efter samme principper, som gjaldt for RO's ord med lodret streg. Det vil sige, at nye sammensætninger lagres, hvis sidsteleddet allerede kendes. Giver ordbogen selv oplysninger om bøjning ved (usammensatte) opslagsord, som ikke kan genkendes på grundlag af Ordbanken, lagres ordet med disse bøjningsoplysninger.

Supplering fra korpus

På dette tidspunkt rummer Ordbanken alle genkendte, forskellige opslagsord fra de fire ordbogskilder og er nu klar til at blive konfronteret med korpuset. I første omgang vil vi gerne vide, hvilke korpusord der er kendte FORMer af kendte LEMMAer. Det undersøges, om det enkelte korpusord findes blandt de former, der kan genereres af HETEROGRAfernes bøjningsoplysninger (BOJNING). I så fald lagres korpusordet som FORM i relation til det eller de LEMMAer, den eller de aktuelle HETEROGRAfer tilhører.

For at kunne foretage denne bestemmelse må vi kunne fortolke lemmaernes bøjningsoplysninger, og hér støder vi på et problem, som vi har arvet primært fra RO. Langt den overvejende del af disse oplysninger stammer fra RO, dog suppleret med oplysninger fra B&H og V&B. Bøjningsoplysningerne er ikke angivet udtømmende, men peger hver især på et bøjningsmønster.

RO er skrevet med danske sprogbrugere for øje, men ikke alle danske sprogbrugere. I forbindelse med præsentationen af den nye Retskrivningsordbog i 1986 blev det sagt, at det "er - og har altid været - en stiltiende forudsætning at de officielle retskrivningsordbøger med tilhørende retskrivningsvejledninger ikke sigter mod de danske skriftsprogsbrugere i bred almindelighed, men mod skriftsprogsbrugere over et vist niveau. Groft sagt er Retskrivningsordbogen først og fremmest beregnet på dem der i forvejen sidder inde med en høj grad af skriftsprogsbeherskelse, og de emner der behandles i Retskrivningsreglerne er først og fremmest dem hvor 'ordentlige sprogbrugere' kan tillade sig at være i tvivl" (Galberg Jacobsen 1987: 33).

Nu er *brugeren* i denne sammenhæng ikke noget menneske, men derimod en række computerprogrammer. Disse programmer indeholder ikke fra første færd den viden, man altså forudsætter af brugerne. Når man automatisk skal tolke RO som udsagn om

normen, rejser der sig derfor det konkrete spørgsmål: Hvad er det disse skriftsprogsbrugere véd, og som ikke står i RO, men som man skal implementere i programmerne for at sikre, at de kan administrere normen forsvarligt?

Opgaven, vi står overfor, er altså at *efterkonstruere* den kompetence, som disse sprogbrugere forventes at sidde inde med. Ellers kan vi ikke få et program til at afgøre, om et vilkårligt tekstord fra korpuset falder inden eller uden for normens rammer. Og kan vi ikke det, kan vi heller ikke leve op til målsætningen om at bruge Ordbanken til at sammenligne norm og usus.

De oplysninger, RO giver om et ords bøjning, peger på et bøjningsmønster, hvor sprogbrugeren - eller her programmet - selv må fylde hullerne ud. Men et program kan ikke bringes til at se, om nogle af de herved dannede former er umulige. Generelt er det fx således, at der ved substantiver er "givet oplysning om bestemt form ental og ubestemt form flertal, når disse kendes. Der er kun givet oplysning om bestemt form flertal, når den ikke dannes ved tilføjelse af endelsen *-ne* til den ubestemte flertalsform" (RO: 18). Ved "*tilfælde* -t, flt. *tilfælde*" er der ikke oplyst om bestemt form flertal, som derfor kan dannes ved at tilføje *-ne* til *tilfælde*: *tilfældene*. Ved "*gentleman*, en, flt. *gentlemen*" er formen ligeledes ikke anført og skulle således kunne dannes efter samme princip, hvilket de fleste danske sprogbrugere nok vil undlade.

En holdning til denne type overgenerering kunne være at sige, at hvis en dannet form er umulig, vil den næppe heller forekomme i korpuset og dermed ikke forårsage støj i form af utilsigtede bestemmelser af korpusord. Imidlertid peger forholdet på en mangel ved denne traditionelle måde at oplyse om et ords bøjning på. En udeladt oplysning om en bøjningsform kan betyde ét af to: enten at formen ikke findes, eller at det er forudsat viden, om den dannes. En sprogbruger "der i forvejen sidder inde med en høj grad af skriftsprogsbeherskelse" vil måske nok kunne skelne, men det kan et program ikke.

Som udgangspunkt har vi valgt at danne netop de former, som bøjningsoplysningerne giver os grund til at antage er mulige. Herunder inddrager vi også de retningslinjer som ellers findes i RO (RO: 13-23; 549-571). Desværre medfører denne normlydige adfærd, at vi kommer til at undergenerere, hvilket direkte har konsekvenser for redaktionsarbejdet. Problemet med undergenereringen er i praksis, at søgninger i korpuset bliver ufuldstændige og dermed upålidelige. Hvis man fx søger i korpuset efter forekomster af adjektivet *smuk* og alene følger RO's mønster "*smuk* to., -t, -ke", finder man belæg på formerne *smuk*, *smukt*, *smukke* og *smukkes*. Man finder ikke i samme søgning *smukkere*, *smukkest* og *smukkeste*, som også er belagt. Vi kan altså ikke beskrive korpuset tilfredsstillende, hvis vi nøjes med at holde os inden for normens eksplicit angivne rammer.

Undergenereringen rammer især - som antydnet - én gruppe af former, nemlig adjektivernes gradbøjning. RO oplyser kun om gradbøjning ved enkelte adjektiver, nemlig "hvor der sker vokalskifte eller ændring til en anden stamme, og ved tillægsord, der ender på *-som*" (RO: 20). Adjektiver, som ikke opfylder disse krav, men som alligevel kan gradbøjes, savner oplysning herom. Det er sagtens muligt at implementere reglerne for, hvordan formerne dannes i et program, men programmet vil i så fald gradbøje alle de adjektiver, det kommer i nærheden af, og ikke kun dem, der faktisk kan gradbøjes.

Når vi ikke i Ordbanken har tilstrækkelig viden om gradbøjningen til generelt at danne formerne, må vi lede efter dem i korpuset. Det gør vi ved hjælp af morfologisk analyse, hvor vi genkender et korpusord som enten komparativ eller superlativ, hvorefter oplysningen om gradbøjning indføres i banken ved det aktuelle adjektiv. Herved får vi i det mindste viden om, hvilke adjektiver sprogbrugerne gradbøjer.

De tekstord, vi ikke har kunnet bestemme på grundlag af bankens aktuelle oplysninger, udgøres primært af sammensætninger, afledninger og bøjningsformer heraf. Hvis vi kan opløse dem, vha. af de orddele vi allerede kender, indføres de i banken. Dette er for sammensætningernes vedkommende ikke helt så enkelt, som det lyder. For det første gør de samme problemer med homografe sidsteled sig gældende, som de gjorde i ordbogskilderne. Vi har altså stadig brug for manuelt at entydiggøre disse sammensætninger. For det andet kan en del sammensætninger tilskrives mere end én analysestruktur. I tilfælde af flere strukturer er det altså nødvendigt med en manuel indgriben for at vælge blandt resultaterne. Fx vil en dansk sprogbruger nok ikke spekulere over, at korpusordet *reklamebudskaber* kan tolkes som andet end ubestemt form flertal af et *reklamebudskab*. For et program, der intet kender til virkelighedens verden, kan ordet imidlertid også tolkes som en *reklamebudskaber* (jf. fx Brandt 1993).

Nu står vi tilbage med en samling korpusord, der hverken er genkendelige bøjningsformer eller sammensætninger og/eller afledninger. Som man kunne vente, indeholder restmængden en del regulære slå- og stavfejl, der er også ord fra andre sprog, ad hoc-dannede interjektioner og nogle helt nye lemmaer, men hvad mere interessant er, så rummer listen også en del *regelmæssige variationer* over de bøjningsoplysninger, vi allerede har i banken samt en del supplerende former. Fx kan flertalsformer af substantiver som if. RO ikke kendes i flertal dukke op hér, fx *polterabender* ("*polterabend*, en." (RO)) og *sugeevner* ("*sugeevne* -n." (RO)). I så fald opdateres Ordbankens bøjningsoplysninger (BOJNING) med oplysning om, at denne brug forekommer i korpuset.

Endelig finder vi her eksempler på afvigelser, som kolliderer med autoriserede former, men som er dannet regelmæssigt, det vil sige dannet ved hjælp af kendte fleksiver og kendte morfografematiske ændringer, fx *brudet* ("*3. brud* -det, brud -dene." (RO)), *sweatre* ("*sweater* -en, -e sweaterne." (RO)) og *karnevallet* ("*karneval* -et, -er" (RO)). Disse afvigelser indføres også i banken, fordi de spiller en rolle, når et helt lemma skal findes i én søgning, men det oplyses at formen står i opposition til normen. Oplysninger om regelmæssig afvigelse er væsentlige, når vi skal undersøge, hvordan sprogbrugerne bruger ordene i virkeligheden.

Anvendelser

Oplysningerne om ortografiske varianter og bøjningsparadigmer bruges af Den Danske Ordbogs konkordansprogrammel, hvorved søgning på én grundform automatisk kan udvides til søgning efter samtlige stovevarianter i alle bøjningsformer (autoriserede eller ej). Desuden bruges bøjningsoplysningerne til første trin af en automatisk homografseparation (ordklassetagging af korpus), og til undersøgelse af, hvilke bøjningsformer af

ordene der faktisk forekommer i korpus; herved fås et empirisk grundlag for Den Danske Ordbogs behandling af bøjningsoplysninger. Ud fra databasen konstrueres endelig de "artikelskeletter", som ordbogsredaktørerne kan benytte som udgangspunkt for redigering af Den Danske Ordbogs artikler. Ingen oplysning fra Ordbanken bliver altså viderebragt i ordbogen, uden at den har været vurderet af mennesker.

Referencer

- B&H: Blinkenberg, Andreas og Poul Høybye: *Dansk-fransk ordbog / Dictionnaire Danais-Français*. 4. udg. ved Jens Rasmussen & al. Bind 1-2. København 1991.
- Brandt, Søren (1993): *Knasegrynteranalyseprogrammer*, Weekendavisens bogtillæg, Bøger, Nr. 17-1993: 13.
- DANLEX 1987a: The DANLEX Group (Ebba Hjorth, Bodil Nistrup Madsen, Ole Norling-Christensen, Jane Rosenkilde Jacobsen, Hanne Ruus): *Descriptive Tools for Electronic Processing of Dictionary Data*. Studies in Computational Lexicography, Niemeyer, Tübingen 1987.
- DANLEX 1987b: DANLEXgruppen (Gert Engel, Ebba Hjorth, Bodil Nistrup Madsen, Ole Norling-Christensen, Jane Rosenkilde Jacobsen, Hanne Ruus; red.): *Ordbøger i Danmark. En oversigt*. København 1987.
- DANLEX 1989: DANLEXgruppen (Ebba Hjorth, Bodil Nistrup Madsen, Ole Norling-Christensen, Hanne Ruus, Jørgen Schack, Peter Widell): *Behov og Brugerprofiler for Dansk Ordbank*. København 1989.
- Dansk Ordbog 1989: Ebba Hjorth, Iver Kjær, Ole Norling-Christensen og Hanne Ruus: *Dansk Ordbog. Plan for en ordbog over dansk Nutidssprog*. 18. december 1989.
- Galberg Jacobsen, Henrik (1987): *Retskrivningsreglerne*, i Selskab for Nordisk Filologi, Årsberetning 1985-1986: 32-35. København 1987.
- Kristensen, Kjeld (1994): *Den Danske Ordbogs tekstkorpus og spORDhunde* (i denne bog).
- ODS: *Ordbog over det Danske Sprog*. Udgivet af Det Danske Sprog- og Litteraturselskab. Bind 1-28. København 1918-56.
- RO: *Retskrivningsordbogen*. Udgivet af Dansk Sprognævn. København 1986.
- V&B: Vinterberg, Hermann og C.A. Bodelsen: *Dansk-engelsk Ordbog*. 3. udg. ved Viggo Hjørnager Pedersen. København 1990.

ODS-Supplementet

Præsentation af bind 1: A(Aa)-Bh. 1992

Sv. Eegholm-Pedersen

Vi plejer at regne Ordbog over det danske Sprog (ODS) til de store nationalordbøger i Vest- og Nordeuropa. Men når vi nævner vores "store ordbog" sammen med de andre fire store det her drejer sig om, nævnt i aldersrækkefølge: Grimm: Deutsches Wörterbuch, Woordenboek der Nederlandsche Taal, The Oxford English Dictionary (OED), Svenska Akademiens Ordbok (SAOB), så er det rimeligt at sætte den beskedent til sidst eller efter en tankestreg som Hans Jonsson i foredraget om de store ordbøgers særart på Oslo-kongressen 1991.¹ Den omfatter med sit 1700-års startpunkt et kortere tidsrum end de andre, der alle går ned til eller (OED) over middelaldergrænsen, og den er i det hele af noget mindre omfang.

Konstituerende for disse store citatbaserede ordbøger er det historiske anlæg. De har to dimensioner, den diakroniske og den synkroniske, således at forstå at de - med Hans Jonssons ord - indeholder en sambeskrivelse af en vis afgrænset tidligere periodes sprog og samtidssproget. De er alle projekter med en meget vid tidshorisont for udgivelsen, og det ligger i sagens natur at beskrivelsen af samtidssproget i ordbogens første partier på en række punkter vil være forældet ved værkets afslutning.

Det altså uomgængelige revisions- og suppleringsbehov er løst eller ved at blive løst på forskellig måde ved de tre af de fem som i dag er fuldført. OED som kom først i hus (1928 efter en udgivelsesfase på 44 år) har også præsteret den tilsyneladende optimale løsning: først et supplement 1933, derefter et større ajourførende supplement (1972-86) og endelig en integreret udgave af hovedværk og supplement, markedsført i trykt og elektronisk form som OED, 2. udgave (1989). Grimm stod med et problem i værkets meget inhomogene karakter. Den meget lange udgivelsesperiode og den noget løse organisation medførte betydelige forskelle i omfang og stoffets behandling i de enkelte dele. Her kunne der ikke lappes med supplement. Løsningen måtte nødvendigvis blive en ny udgave, en Neubearbeitung. Hvis første etape A-F er igang (der er foreløbig udgivet 1/3 af bogstavet A og bogstaverne D + halvdelen af E).

Hvad nu ODS angår, så var udgivelsesperioden ganske vist relativt kort (1918-54), men dog lang nok til at en ajourføring af ordstoffet i de tidligere dele følte meget påkrævet ved værkets afslutning. Og hertil kom at også ODS havde et homogenitetsproblem. Langtfra på samme måde som Grimm, men dog alligevel. Måler man den rent kvantitative behandling af de enkelte bogstaver i ODS med gennemsnitstal for andre danske ordbøger, vil man finde en betydelig skævvridning i den store ordbog. Bogsta-

verne A - F fylder gennemsnitlig ca. 26% af totalomfanget i ordbøgerne, i ODS kun 17. Til gengæld er slutningen af alfabetet, bogstaverne S - Ø som gennemsnitlig lægger beslag på 29%, i ODS svulmet op til 39%.

Det er klart at for brugeren betyder denne skævhed - som har sin baggrund først og fremmest i det på flere punkter helt utilstrækkelige materiale ved redaktionens begyndelse - den betyder at de oplysninger man får om ord og ordbrug i begyndelsen af alfabetet er langt fra så fyldige og tilfredsstillende som dem der kan hentes i de senere bind. Og det må blive en vigtig opgave for et supplement at søge i nogen grad at udligne denne forskelsbehandling. Lis Jacobsen, ordbogens navnkundige administrator, formulerede allerede i 1950'erne et ønsket supplements opgave som den dobbelte: ajourføring og egalisering. Når vi i dag hvor første bind af det planlagte 5-binds supplement foreligger taler om at Supplementet skal afrunde ODS, ligger der altså heri de to ting: dels en tilføjelse af det nye stof, nye ord og ordanvendelser fra redaktionsperioden, men vel at mærke: ikke ud over denne. Dels en generel styrkelse af ordbogens svagere dele, specielt de første 5-6 bind.

Kilderne.

I det materiale som ODS-Supplementet (ODS-S) bygger på indgår en oprindelig stamme af citater og henvisninger fra den gamle ODS-redaktionsperiode. Den var på 350.000 sedler, men af en noget tilfældig karakter. ODS havde ikke som SAOB et fast excerpistkorps. Redaktørerne excerpereede efter tid og lejlighed, men denne excerpering var først og fremmest rettet mod de under arbejde værende bogstaver eller de ikke redigerede afsnit.

Den målrettede excerpering som over en årrække blev gennemført med henblik på et supplement, skulle altså dels søge at indfange det nye ordstof fra udgivelsesperioden, dels bøde på erkendte mangler i det gamle materiale. Det drejer sig her især om 1700-tallet, hvor navnlig den midterste del var et forsømt område, både mht. fag- og skønlitteratur, og om fremmedordene generelt. Der er i indledningen til ODS-S I givet en redegørelse for denne supplerings, og jeg skal her blot præcisere et par punkter.

Det samlede materialegrundlag er i dag en seddelsamling på ca. 1.3 mill. sedler. Dette tal skal sammenholdes med det anslåede tal for den gamle ODS som var på 2 1/2 mill. Dvs. at der som grundlag for hvert af de 5 supplementbind gennemsnitlig ligger et betydeligt større materiale end til hvert af de 27 tekstbind som hovedværket omfatter. Det skulle give en nogenlunde fast grund at træde på, men også et meget større materiale at bearbejde til de enkelte bind.

Skal man sige lidt om de forskellige kildetyper og deres indbyrdes vægtning, støder man på den vanskelighed at der ikke foreligger statistisk materiale med oplysning om excerptmængder fra den første del af indsamlingsperioden. Men med dette forbehold kan man nævne følgende som de vigtigste genrer:

Først *aviser* - de tegner sig alene for 17-18% af den halve mill.-tilvækst i den sidste indsamlingsfase efter 1965. Som bekendt er avissproget en fremragende kilde for den

der vil følge den aktuelle sprogbrug og finde tidlige belæg på nydannelser. Det har da også fra første færd været et ønske at udbygge og også styrke ODS generelt på dette punkt. Avisexcerperingen har været både omfattende og alsidig. Hovedstadspressen i 1900-tallet, specielt perioden 1915-55, udgør hovedstammen, men også de to foregående århundreders presse er medtaget, og også provinspressen er inddraget for at få så bred en dækning som muligt. Også ugeblade og enkelte tegneserier er repræsenteret i samlingen, dog med et mindre antal excerperter.

Dernæst *skønlitteraturen*, som var den absolut dominerende kildetype i det oprindelige ODS-materiale og er et tyngdepunkt i ordbogen. ODS er ubestrideligt en smuk litteratur- eller klassikerordbog med fyldig repræsentation af periodens hovedforfattere. Excerperingen til Supplementet - 11-12% af tilvæksten efter 1965 - har først og fremmest været rettet mod hovedforfatterne i ODS' udgivelsesperiode, 20'ernes, 30'ernes og 40'ernes forfattere, alle repræsenteret med en række værker og fyldige excerptmængder. Hvad de ældre forfatterskaber angår har nye udgaver af breve og dagbøger givet materiale til samlingen. Men især vigtig var en genlæsning af en række klassiske forfattere, specielt med henblik på fremmedordene, som var stærkt underrepræsenteret i det oprindelige materiale.

Andre højt prioriterede kildetyper er *tidsskrifter*, især almenkulturelle eller politiske, litterære, men også nogle rent faglige tidsskrifter har leveret stof fra ældre og yngre tid. Videre *encyklopædier*, fra det 20. århundrede og tillige nogle vigtige fra 1800-tallet, samt officielle tryk (love og forordninger m.m.). Hvad sagprosaen iøvrigt angår er der excerperet indenfor de forskellige områder som kulturhistorie, topografi, politik, nyere historie og faggrupper som jura, medicin, økonomi, teknik, håndværk og nye gebeter som fjernsyn og film.

Det bør nævnes at vi opfatter og udnytter Dansk Sprognævns omfattende samlinger som et yderst værdifuldt baggrunds- og støttemateriale. Supplementet og Sprognævnet mødes på grænsen 1955, slutpunktet for ODS-S og begyndelsesåret for sprognævnsregistreringen. Ved ord og ordanvendelser i nærheden af 1955-grænsen er det af stor betydning at have et udvidet bedømmelsesgrundlag i foreliggende sprognævnsцитater. Og Supplementets tidsgrænse er elastisk i den forstand at citater efter 1955 godt kan medtages og lejlighedsvis medtages i artiklerne hvis vi ikke synes ordene er rimeligt belyst ellers.

Supplementet bygger således på en stor - eller skal vi sige: ganske stor - citatsamling, altså en traditionel excerptsamling. I dagens leksikografiske situation er de databaserede tekstkorpuser med svulmende millional for registrerede ordforekomster ved at være standardudstyr, evt. som supplement til foreliggende excerptsamlinger.² Supplementets materialeindsamling var imidlertid jo væsentlig afsluttet på et tidspunkt hvor etableringen og håndteringen af databaserede tekstkorpuser endnu ikke var tilstrækkelig gennemprøvet. Og stillet overfor den nye tekniks fristende muligheder må vi nu hvor vi er langt inde i produktionsfasen med deadlines og tidspres, konstatere at der er lukket af for tilgangen af nyt materiale. Det er ikke muligt nu at anlægge nye korpusstrategier. Hertil kan så føjes den betragtning at traditionel excerpering dog beholder sin værdi ved

specialopgaver som ODS-S ligesom fx. ved nyordsregistreringen i Dansk Sprognævn. Den selektive excerpering, foretaget af kyndige excerpister og så at sige med fingeren i ODS har givet os et - i dobbelt forstand - udsøgt materiale hvor det relevante stof ikke drukner i totalmængden af citater.³

Indholdet: tilføjelser og rettelser.

Supplementets tidsramme er som allerede berørt den samme som hovedværkets, nemlig perioden fra ca. 1700 til 1955. Det er vigtigt at sige eller gentage at vi går principielt ikke ud over dette tidsrum og supplerer indenfor dette.

Et hovedanliggende for Supplementet er ifølge sagens natur det nye sprogstof, nye ord og ordanvendelser fra 1900-tallet, det som er kommet til efter de enkelte binds redaktionsår. Generelt kan det siges at nyordsstoffet fra denne periode er præget af fremmedord, herunder internationale betegnelser indenfor forskellige fagområder der er blevet så almindelige i fællessproget at de er medtaget (*Allergi, Ambivalens, Aspirin, Bakelit* osv.), samt af sammensætninger og afledninger både indenfor det fremmede og hjemlige ordstof.

Sammensætningerne udgør jo som bekendt i danske og nordiske ordbøger den absolut dominerende ordtype. I ODS efter stikprøver⁴ over tre fjerdedele af alle opslagsord. Også i supplementspalterne vrimler sammensætningerne, og det nye ordstof har i en række tilfælde givet en kraftig forøgelse af bestanden i forhold til den gamle ODS. Ordet *Atom* er gammelt i sproget, men bliver med den tekniske udvikling i 1900-tallet produktivt som 1. led med forskellige betydninger i sammensatte ord, og Supplementet føjer til den gamle sammensætningsløse ODS-artikel et par spalter med et par snese nye sammensætninger fra redaktionsperioden: *-affald, -alder, -bombe, -energi, -kraft, -spaltning* osv. Endvidere med eksempler - i en samleartikel - på brugen i jargon eller slang om noget voldsomt eller imponerende: øjeblikdannelser som *Atomsucces* (tegneserie 1947), eller *Atommanden* født i et lykkehjul (sportsreportage 1952). Andre eksempler på meget lange og i forhold til ODS markant forøgede sammensætningsrækker er *Arbejder-* (med 64 +-artikler til ODS' 15 og *Arbejds-* (med 204 +-artikler i forhold til de foreliggende 152). I de to sidste tilfælde går eksempel-rækken tilbage før år 1900 og afspejler interessant den tekniske og samfundsmæssige udvikling fra slutningen af 1800- til midten af 1900-tallet. Det er vigtigt at dette stof tilgodeses bredt i de store ordbøger, og man tør vist sige at Supplementet har været generøs på dette punkt. Som prøve på supplementspalterne er der til sidst i denne artikel vist en ny sammensætningsrække (*Banan-*) med eksempler både på nye tekniske og andre faglige ord og på hverdagsgloser og idiommer (*glide i en bananskal*) som blev almindelige i 1900-tallet efter at bananerne var blevet en del af det almindelige kulturbillede og forbrugsmønstre. Den tilsvarende ODS-artikel - der er ingen sammensætninger - fylder kun 9 linier.

En betydelig del af de nye supplementord er ældre i sproget end det pågældende ODS-bind og har i flere tilfælde også været repræsenteret i ordbogens oprindelige samlinger. Der foreligger her for supplementredaktionen et oplysende materiale i de lister

som til intern brug blev ført i ODS-redaktionen over ord som var til overvejelse mht. medtagelse/ikke-medtagelse. Når vi kan konstatere - for de første binds vedkommende - at flertallet af de nedstemte og altså adgangsformentede ord fra disse lister får plads i Supplementet, så behøver det ikke at betyde at optagelseskriterierne er afgørende forskellige i ODS og Supplementet. Der kan være tale om at ordenes status har ændret sig, ord der har været følt som specielt faglige kan senere have fået en større udbredelse. Rammerne er dog udvidet lidt, og specielt på ét vigtigt punkt er Supplementets praksis nødvendigvis mere liberal end ODS', nemlig mht. fremmedordene. Den mangelfulde dækning af dette stof i den gamle ordbog har ofte og med rette været kritiseret. Der er her tale om ikke blot de fagligt eller teknisk betonedede internationale ord, men det almindelige fremmedordsstof som udgør og i ODS-perioden har udgjort en del af det litterære ordforråd og det almindelige talesprog. Baggrunden for mangelsposterne var som nævnt den oprindelige materialefattigdom. Senere blev materialet bedre, men den forsigtige linie der var lagt mht. medtagelse af fremmedordene ved projektets start, kom i nogen grad til at præge hele værket. Hvor meget der mangler af det ældre stof og her først og fremmest de romanske fremmedord, kan man se fx. ved en sammenligning med det materiale der er fremlagt i Holbergordbogens 5 bind. Nu kan vi altså udfylde en række lakuner både mht. ord som var almindelige i 1700- og 1800-tallet, men sjældne i det 20. århundrede: *abandonnere, acquittere, admirabel, admittere, allegere, Attention, Bassesse*. Og ord som er udbredt (også) i 1900-tallet: *abrupt, absorbere, adækvat, Aggression, ambitios, Aspekt*.

Supplementstoffet fremtræder dels som nye artikler, dels som tilføjelser (sjældnere: rettelser) til eksisterende artikler. De ialt ca. 10.350 artikler i bind I fordeler sig nogenlunde ligeligt på de to kategorier.

Tilføjelserne kan gælde nye betydninger, især naturligvis fra det 20. århundrede. Substantivet *Baand* får med den tekniske udvikling nye anvendelser om lydbånd til båndoptagelser og indenfor radiofonien om bølgeområde, frekvensbånd, desuden om transportbånd, "endeløst bånd". Ordet *Aand* fik en ny anvendelse som skældsord: tåbe, idiot, hvortil adj. *aandet*).

De faste forbindelser (kollokationerne) får som helhed en fyldig behandling i ODS, dog også her med forbehold for de første bind hvor fraseologien er mere tilfældigt belyst. Supplementet kan her føje meget nyt - ældre og yngre - stof til, ofte i form af henvisninger til en bestemt frases behandling i de senere dele af værket. I ODSS-artiklen *Ben*, der føjer godt 4 spalter til ODS' 9, er der mest tale om forbedringer af det fraseologiske stof, med ialt en 35 nyregistrerede faste forbindelser og idiomer: *gnave ben* (snorke), *lette (på) benet, få en til det andet ben, stable på benene* osv.

Meget ofte består tilføjelsen i tilførelse af nyt citatmateriale. Det svage grundlag i begrundelsen af udgivelsesfasen viser sig i mange artikler med citatfattigdom sammenlignet med senere ODS-standard, undertiden i den ekstreme grad at der slet intet belægsmateriale bringes (småartikler som *Absolutist, Absurditet, Admiralinde, Adoptivbarn* osv.).

De mange tilfælde af tilbage- og fremdateringer som Supplementet giver, er nærmere

analyseret af Anne Duekilde andetsteds i denne bog.

Man kan sige at supplementredaktionen i bind I har arbejdet på en taknemmelig baggrund. Det er ikke svært med den udvidede materialebasis både at udfylde lemmalakuner og at forbedre en række af de artikler der har behov herfor. Som regel er det heller ikke problematisk at indføje det nye stof i de gamle artikler. Det forekommer sjældent, men det forekommer dog at en ODS-artikel er erstattet af en nyskreven, se mindre artikler som *Afspænding, Agitation, Barbermaskine*.

Det forekommer også - det må tilføjes - at man i og for sig erkender et supplerings- eller revisionsbehov, men ikke kan gøre noget ved det. Her må nævnes de meget store og svære artikler, typisk en række formordsartikler. Der er en betydelig forskel mellem præp.artiklerne i de første og de senere bind af ODS, sml. fx. *af* med *pa*, *fra* med *til*. Og den behandling som de store pron.artikler under D har fået, føles knap tilfredsstillende i dag. Men de forbedringer som kan finde sted i disse artikler vil være af beskeden og delvis tilfældig art. Supplementet kan ikke her bringe behandlingen op til senere ODS-standard.

Hvad angår egentlige rettelser til ODS-artikler, så kan man konstatere - med en kompliment til de gamle redaktører og medarbejdere - at vaskeægte trykfejl, forkerte kildeangivelser m.m. er forholdsvis sjældne. Ordbogen har aldrig meddelt trykfejlslistor, men en række notater om fejl er naturligvis gennem årene indkommet og vil indgå i Supplementet. Korrigerende indgreb i den redaktionelle tekst vil selvsagt forekomme, dels i tilfælde hvor der er tale om egentlige fejl eller inkurier mht. citatolokning eller -placering, dels og især i de uundgåelige tilfælde hvor den gamle redaktør ikke har været godt nok underrettet, har formet en definition eller en brugsbestemmelse på et for spinkelt grundlag, som Supplementsamlingen nu kan forbedre.

Når Supplementet engang foreligger færdigt med sine 5 bind vil det store stof, som der her er givet nogle prøver på, fordele sig ulige over bogstavrækken. Det foreliggende 1. bind, der går fra A til Bh, svarer til 1 1/2 ODS-bind og giver altså så høj en suppleringsprocent som 67. De følgende to bind, der bliver udsendt i 1994 og 96, bringer Supplementet frem til og med bogstav F og svarer til 4 ODS-bind, dvs. en suppleringsprocent på gennemsnitlig 50 (i en faldende kurve). For resten af alfabetet regner redaktionen med en yderligere faldende kurve ned til procenter på ca. 10-15 (P-R) og 3-5 (S-Ø). Samlet får vi altså og måtte nødvendigvis få et supplement med en betydelig forvægt - grafisk et spejlbillede af hovedværket.

Problemer.

Jeg skal runde af med kort at berøre et par generelle og specielle problemer i forbindelse med supplementarbejdet. At udarbejde et sådant supplement er en meget bunden opgave i den forstand at den jo er styret af og forudsætter en accept af den gamle ODS-struktur både på makro- og mikroplanet. Der kan justeres lidt, men der kan ikke ændres afgørende. Supplementet skal hele tiden fungere sammen med og læses sammen med hovedværket.

En iøjnefaldende ulempe ved den gamle ODS som Supplementet må acceptere, er de i dag på nogle punkter forældede lemmaformer. ODS er i sit redaktionssprog avanceret for sin tid idet der bruges små bogstaver i substantiverne (men dog ikke bolle-å). Men opslagsformerne følger naturligvis den gamle retskrivning før 1948 med store begyndelsesbogstaver i substantiverne og dobbeltskrevet aa som betegnelse for å. Suppl. må her følge trop og altså alfabetisere å-ordene som a + a (*båd* før *bad* osv.), hvad der jo virker ligeså forkert som den gamle indordning af æ og ø som a + e og o + e, som man finder i ordbøger fra 1700- og begyndelsen af 1800-tallet.

Artikelstrukturen er principielt den samme i Supplementet som i hovedværket. Men den er nu formaliseret og derved på nogle punkter præcisere og mere konsistent idet manuskriptet løbende bliver indskrevet i en feltdatabase. Der er dog én vigtig forskel: de nye artikler i Supplementet har ingen udtaleangivelse, og der meddeles ikke rettelse til evt. forældede angivelser i ODS (som fx. når ordet *Agurk* i bind I har udtaleangivelse med o). Ændringen er en nødvendig konsekvens af tidsafstanden mellem redaktion og den behandlede periode: vi kan jo ikke i 90'erne sidde og statuere normaludtaler 40 år tilbage.

Et særligt problem udgør brugsbestemmelserne. ODS opererer med et nuanceret og fint gradueret system af markører til angivelse af brugsområde og brugelighed. Alle ord som ikke på redaktionstidspunktet var fuldt brugelige i neutralt rigssprog (det almindelige skrift- og talesprog) skulle forsynes med vejledning herom, altså angivelse af det særlige brugsområde eller stillag, eller nærmere bestemmelse af frekvens eller begrænsning til en bestemt periode. Systemet blev hele ordbogen igennem administreret særdeles samvittighedsfuldt. Der er rundt omkring i artiklerne mange omhyggelige, undertiden meget detaljerede brugsangivelser, som er af stor værdi for ordhistorien.

Det er klart at Supplementet som hovedværket skal give besked om ordenes status, både de nye ords status og (i et vist omfang) eventuelle ændringer i det gamle ordforråds frekvens og brugsområde. Men det er ligeså klart at supplementredaktørerne her arbejder under andre betingelser end de gamle ODS-redaktører. Vi kan ikke som de overfor ordstoffet fra 1900-tallets første halvdel give en samtidsbedømmelse. Selvom begge de nuværende redaktører er på den i denne forbindelse helt rigtige side af de 60, med (evt. også sproglige) erindringer fra 30'erne og 40'erne, så lader det sig ikke gøre på fri hånd, uden fyldig materialebaggrund at decidere med sikkerhed mht. de finere nuancer i den sproglige usus, frekvensforhold osv.

Konsekvensen for Supplementet og dets forhold til hovedværket er en større tilbageholdenhed og forsigtighed på dette felt, brugsbetegnelserne bliver mindre udførlige, og så således at det ubetegnede område af sprogbrugen, kredsen af ord der ikke forsynes med nogen brugsanvisning, udvides noget. I ODS finder man meget hyppigt bogtegnet som markør for ord der kun eller fortrinsvis bruges i skriftsprog eller litterært påvirket talesprog. Dette tegn, som i Supplementet er erstattet af ordet skriftspr., bruges væsentlig sjældnere her.

Hvad endelig angår det rent ordbogstekniske problem om sammenkobling af hovedværk og Supplement, dvs. tilordning af supplementartiklerne til ODS-artiklerne, så har

ODS-S valgt den nærliggende løsning overalt tydeligt at markere tilføjelser og nyt stof med et plus-tegn. Nye artikler indledes af et +, og tilføjelser til gamle artikler, der kan sætte ind på alle punkter i artikelstrukturen, indføres også med et plus.

Supplementet er med sine mange plus'er og spalte-liniehenvisninger skræddersyet til sit hovedværk og - forhåbentlig - således indrettet at det er relativt problemløst for brugeren at gå frem og tilbage mellem de to tekster. Det ville selvfølgelig være nemmere og mere brugervenligt at have oplysningerne samlet, i samme bind. Jeg har nævnt Oxfordordbogens smukke tretrinraket: hovedværk - supplement - integreret udgave. Men denne model, hvis sidste fase for øvrigt er langt fra problemløs således om den er realiseret i den store engelske ordbogs "2. udgave",⁵ den kan endnu vanskeligere overføres til den store danske ordbog, fordi ODS er redigeret som den nu engang er: med megen og vel overvejende vægt på samtidskomponenten i forhold til den historiske komponent, altså med vægten på den nøjagtige og minutiøse redegørelse for redaktionstidens sprog. Det er vanskeligt at se hvordan man skulle håndtere de mange tidsbundne oplysninger om ordenes status osv. i en integreret udgave hvor hele stoffet ansues ud fra en historisk synsvinkel. Der måtte da blive tale om en væsentlig omarbejdet ODS.⁶

Her og nu drejer det sig selvfølgelig om at få de planlagte 5 supplementbind vel i hus. Får vi det vil der i hvert fald foreligge en afgørende forbedret ODS, men altså nu ikke en samtidsordbog med historisk komponent, men en historisk ordbog over dansk rigssprog 1700-1955 (dog med hovedvægt på det sidste halve århundrede).

Noter

1. R.V. Fjeld (udg.): *Nordiske Studier i Leksikografi*. Oslo 1992. 231-34.
2. Se fx. rapporten fra den islandske ordbog, Nord. Stud. Leks. 390-402.
3. Rapporten fra den islandske ordbog understreger i sin konklusion værdien af den traditionelle excerptsamling også for dette projekt.
4. Aage Hansen: *Moderne Dansk*. Kbh. 1967. I 253.
5. Jf. Noel Osselton (anm.) *The Oxford English Dictionary*. Sec. Ed., i: *International Journal of Lexicography* VI,2. Oxf. 1993. 124-31.
6. Se også Anne Duekilde: *Supplement til Ordbog over det danske Sprog*, i: J. Louis-Jensen & J.H.W. Poulsen (udg.): *The Nordic Languages and Modern Linguistics* 7. Tórshavn 1992. 181-90, spec. 185f.

ODS:

Banan, en. [ba'na:n] fl. -er. (af port., sp. banana; opr. et afrikansk ord) frugten af pisangplanten, *Musa paradisiaca* L. (jf. Adamsfigen). vAph.Nath.II.331. Funke. (1801).II.455. Frugten (af pisang) (der kaldes „Banan“) er et Bær af Form som en glat, but trekantet Agurk. Warm.Frøpl. 162. || & om selve planten. Funke.(1801).II. 455. Sal.II.513.

ODS-S:

Banan, en. + (i ældre spr. ogs. Bannana oign. fl. -s). ældre o.a.: Bannanas eller
 15 Broddi er en Frugt som i Europa haver ingen at lignes ved. ETilleman.Gvinea.(1697). 139. Bannanes. Pflug.DP.752. Figen. Yams. Bananoes. Reiser.II.378. *Alberti, Alberti. han gaar fra Dør til Dør | og sælger 3 Bananer for 25 Ør. Cit.1908.(Spectator.1948/49. 81). 1892 .. De første Bananer kommer til Byen i denne Tid. KraksTidstavler.(1930). 93. Blomsterne (på bananplanten) sidder flere sammen („en Haand“ Bananer). Nytte-
 25 planter.162. *Mæn når der kommer en båd med bananer, så har vi glemt alle sorger og savn. vise 1945.(Vogel-førg.BO.*425). + || (i 1930'erne) spøg. om vegetar. navnlig den ene af (ernæringsspecialisten Mikkel Hind-
 30 hedes) forsøgspersoner .. kaldet „Bananen“, blev berømt .. kapløbet mellem „Bøffen“ og „Bananen“. DaglDanm19og20årh.II.384. + || (slang) om genstande hvis form kan minde om en banan. om cigar: de almindelige
 35 slanke Haandcigarer er „Bananer“. Tidsskr.f. Tobak.1/12 1924.9.sp.2. En almindelig „Banan“ kan ofte være lige så god eller bedre end facon-cigaren. DKøbm.131. om projektil: Bom.S.II.110. om en finger: Da vi sikrede
 40 Danmark.(1915).40. om det mandlige lem: Bom.S.162. Rifbjerg.Sommer.(1974).48. + || som (eufem.) udtr. for hård behandling; kle (jf. bl.a. Kaffe 2.4. Kanel 2.2. Konfekt. Tobak). Du kan stole paa at der er Bananer i
 45 Luften til dig. VBergstrøm.HN.13. jf.: En hård dom, der medfører at man har f.eks. 4 år (4 bananer), hedder et øksehug. Bom.S Ordb.207. + || (slang) om penge. Bom.S Ordb.108. jf.: Være bananstærk 'have penge
 50 på lommen'. Cit.1945.(DSt.1948.119). + **Banan-bøsning**, en. (jf. Bøsning S og -muffe S; tekn.). A Warrern.DaEngTeknO. (1949). + -flue, en. lille flue af slægten *Drosophila*. den i de senere Aars Arvelig-
 55 hedsundersøgelser saa stærkt anvendte Bananflue (har) kun 4 Par Kromosomer. Vore Sygd.I.258. NySal. + -mand, en. gadehandler der sælger bananer fra en banan-

vogn. Bananmanden begynder sin Arbejds-
 dag paa Kongens Nytorv. Riget.23/3 1911.3. sp.5. TandR.A.7. Bananmanden har faaet
 Kammeraten til at gaa alene med Vognen
 5 den Times Tid. HHLund.Lørdag.(1933).89. + -mel, et. (jf. Melbanan; fagl.). Bananmel .. fremstilles i temmelig stor Stil i Mellem- og Sydamerika. Nytteplanter.163. „rigtig“ Mad til Baby – med ægte Bananmel. Beri
 10 Tid.19/6 1945.Aft.3.sp.3. + -modneri, et. 1927.(se u. Modneri). + -modnings-anstalt, en. AvisAarb.1930.2. Lovtid.1955.A. 211. + -muffe, en. (jf. I.Muffe 2.2 og -bøsning S; tekn.). Bananmuffe. Gallalit. fineste
 15 Udførelse. KatalogNr.21. Cyclekompagniet A/S.(1937).27(KglBiblSmåtryk). + -republi- blik, en. (efter eng. banana republic; avis- spr.) nedsæt., om lille stat, især i Mellem- amerika, med en (ustabil, uudviklet) øko-
 20 nomi der er stærkt afhængig af banan- eksport. behandle Guatemala som „banan- republik“. Sal.T.1953-54.525. JyllP.24/6 1954. 4.sp.6. + -skal, en. (jf. -skræl S). Banan- skaller. Levende Billoder fra Sommeren
 25 1906 (af folk der falder ved at træde på ba- nanskaller). Blækspr.1906.45. glide i en bananskal, som udtr. for at noget ikke lykkes. den friske Musik, der forekom lvar at strømme saa purt og naturligt som selve
 30 Skovens Kildevæld, modsat meget af Nuti- dens, som i Pausen mellem to Eksprestog gør sine fortvivlet-originale Saltomtaler og sluttelig glider i en Bananskal. JacPaludan. MM.100. Der skal megen Resignation til for
 35 begge Parter, for ikke at glide i lignende Ba- nanskaller, som Venstre og de Konservative. Finanstidende.1928/29.596. || (jf. sv. banan: sø.-jarg.) om hængekøje. Bom.S.II.146. Hæn- gekøjerne – populært kaldet bananskallerne – vil gradvis blive erstattet med faste køjer.
 40 Beri.Tid.23/7 1960.7.sp.5. + -skræl, en. (jf. -skal S). Norman-Hans.MA.145. man har da heller ikke Mod paa aa sidde aa gnave Bananskræller. naar a' man er ude! FrSkous-
 45 bo.ES.99. + -stik, et. (-stikker. 20 Stk. Ba- nanstikkere. PolitiE.Kosterbl.18/3 1925.2.sp. 2). (jf. ty. bananenstecher, eng. banana plug) enkeltpolet stikprop. 4 Bananstik .. Jordklemme (til en radioantenne). OrdbS.(re-
 50 klametryksag 1926). BeriKonv.II.257. ved hjælp af et par bananstik forbandt (han hæ- reapparatet) med wirerecorderen. Informa- tion.17/7 1954.4.sp.4. + -vogn, en. (jf. -mand S). Banan-Sæsonen er inde. Banan- vognene har atter indtaget deres Plads i det københavnske Gadebillede. Riget.23/3 1911.3. sp.5. Dagens Vær. De ved, om Byen og Pi-
 gerne og Bananvognen. JacPaludan.S.57.

Ekvivalensproblematikk i Nordisk leksikografisk ordbok

Ruth Vatvedt Fjeld

1 Hva er ekvivalens?

Nordisk leksikografisk ordbok (NLO) er et ordboksprosjekt der nordisk terminologi i leksikografien skal samles og defineres. For hver term skal det oppgis ekvivalenter på de nordiske språkene (norsk, dansk, svensk, finsk, islandsk), samt engelsk, tysk og fransk. Lemmautvalget er presentert av Hans Kristian Mikkelsen i denne rapporten. Lemmene i ordboken skal ordnes etter de norske bokmålskvivalentene, da en anser at norsk bokmål er den språkformen som oppleves som den mest fellesnordiske. Det blir derfor av særlig viktighet hva som velges som norske bokmålskvivalenter, siden disse vil få en viss betydning for hva som blir ekvivalent for de øvrige nordiske språkene. Ekvivalentvalget må bestemmes ut fra interne norskspråklige hensyn, men selvfølgelig også ut fra forhold i de øvrige nordiske språkene. Å finne de beste ekvivalentene vil i denne sammenhengen si for hvert lemma å finne den ekvivalenten som

- a) gir best mulig dekning av utgangslemmaet
- b) er mest mulig lettforståelig for nordiske brukere
- c) gir best hjelp ved tilegnelse av internasjonal metaleksikografi
- d) samsvarer best med usus i nordisk leksikografi (deskriptivt)
- e) gir mest mulig enhetlig nordisk terminologi (normativt)

Ekvivalens har opptatt oversettere helt siden man begynte å interessere seg for teoretiske problemer ved å transformere et budskap uttrykt i ett språk til uttrykk i et annet. Det er selvfølgelig fordi det ofte er vanskelig å finne uttrykk i to forskjellige språk som nøyaktig svarer til ett og samme innhold. For eksempel skriver Wilh. von Humboldt i 1796 at det er umulig å finne en helt dekkende ekvivalent i oversettelse. Han sier at enten er man utro mot kilden, eller så er man det mot sin egen nasjon (han forutsetter morsmål som målspråk). Dette er uttrykk for et dilemma som også gjelder i all tospråklig leksikografi.

Innen translatologien er det forskjellige syn på ekvivalensproblemet, det spenner fra tro på absolutt oversettbarhet til det motsatte - at all oversettelse er totalt umulig. I praksis vet vi at en som oftest befinner seg midt imellom. Mye kan oversettes, men noe går også tapt. Werner Koller (1992) betegner relasjonen mellom måltekst og kilde tekst som "ekvivalensrelasjonen". Roman Jakobson kaller det "translation propre" eller "interpretation of verbal signs by means of some other language" (1959:233). "Transem" har

også vært brukt som term (K.M. van Leuven-Zwart 1989:157). Terminogene Picht & Draskau (1985) betegner synonymi og ekvivalens som to svært like fenomener: i terminologi er synonymi identitet mellom begreper uttrykt med forskjellige termer innen ett og samme språk, ekvivalens er det samme uttrykt ved hjelp av to eller flere språk.

Å finne den rette ekvivalenten dreier seg om å kople sammen samsvarende språklige enheter av forskjellig størrelse og type (fra fonem til tekst) i forskjellige språk. Det kan gjøres rent deskriptivt basert på tospråklige informanter, eller det kan oppgis preskriptivt hvordan de språklige enhetene skal koples til hverandre. Det vil alltid måtte være et avveiningsspørsmål hvilke ekvivalenter som skal koples sammen. De to- eller flerspråklige ordbøkene er viktige hjelpemidler for oversettere, men vil aldri være noen "fasit" ved oversettelse av tekster, da kontekstuelle og pragmatiske faktorer også styrer valget av ekvivalenter. Hvor store ytringsenheter som skal oversettes, fra enkeltord til hele tekster, har betydning for hvilke typer ekvivalensproblemer som kan oppstå.

I leksikografi og terminologi er ekvivalensproblemet noe enklere enn i oversettelse av hele tekster. Systematisk utarbeidelse av normert terminologi innenfor et fagområde foregår i to trinn. På det første trinnet i den terminologiske leksikografien må en undersøke fagområdet og det fagvokabularet som faktisk blir brukt - altså et rent deskriptivt arbeid. Det er det som hittil er gjort i NLO i forbindelse med lemmautvalget, å kartlegge hvilke benevnelser som anvendes med hvilket begrepsinnhold i nordisk og internasjonal leksikografi. Neste trinn er så det normerende arbeidet med terminologien, nemlig å bygge opp et terminologisk system og å anbefale noen former framfor andre. Dette arbeidet innebærer det å avgrense nærliggende begreper, deretter å fastlegge et begrepsystem ved vurdering og utvalg av benevnelser som tilordnes begrepene, samtidig som benevnelserne gis definisjoner. Dette må foregå som en prosess der det stadig er vekselvirkning mellom termvalg og definisjon. Til slutt velges hvilke delelementer de nye termene skal bestå av. Det er særlig i det siste punktet at ekvivalensproblemet melder seg med full styrke. Mange valg i dette normerende arbeidet blir styrt allerede idet det velges ekvivalenter for de forskjellige lemmaene, og det kan være en fare at en fastlegger mer enn en egentlig ønsker ved å være for språkbundet ved valg av ekvivalenter.

Med utgangspunkt i en del eksempler fra det foreløpige materialet i NLO skal det her skisseres noen problemer man må ta stilling til i forbindelse med valg av ekvivalenter, og videre drøftes hvilke hensyn som spiller inn ved valget mellom flere alternativer. Prosjektdeltakerne i de forskjellige landene har gitt forslag til ekvivalenter, og det er disse forslagene følgende oversikt bygger på.

2 Valg av ekvivalent

Utgangslemmaet er ofte en internasjonal term, det vil her si engelsk, tysk eller fransk, som skal få norsk, dansk og svensk ekvivalent. Det gis også islandske og finske ekvivalenter, men ekvivalensproblemene her er så mye annerledes enn for de tre andre språkene at de ikke blir tatt opp her. Termer som har latinsk eller gresk opphav, kalles for "internasjonalismer". Gresk og latin fungerer jo nærmest som lingua francae i de fleste

Ruth Vatvedt Fjeld

fagspråk, naturlig nok også innen leksikografi og andre språkvitenskapelige disipliner.

Ekvivalentforslagene kan sorteres i tre typer etter hva slags delemmenter de består av:

| <i>Utgangslemma:</i> | <i>Ekvivalentforslag:</i> |
|---------------------------|---|
| a internasjonalsisme | → internasjonalsisme / nordisme |
| eks: neologisme | → neologisme / nyord |
| biskopal | → biskopal / toveis |
| b fransk/engelsk/tysk | → internasjonalsisme / nordisme |
| eks: teildiomatisk | → partielt idiomatisk / delidiomatisk |
| front matter | → pretekst / fortekst |
| c sms. internasjonalsisme | → internasjonalsisme+internasjonalsisme / internasjonalsisme+nordisme / |
| nordisme+nordisme/ | nordisme+internasjonalsisme |
| eks: regroupment | → regruppering / omgruppering |
| delemmatisation | → delemmatisering / avlemmatisering / lemmaoppløsning |

Teoretisk sett kunne det selvsagt bli mange flere ekvivalenttyper, men det er disse typene som går igjen i materialet.

Ved valg av ekvivalent er det en del avveininger og vurderinger en må ta, styrt av bl a de hensyn som er nevnt innledningsvis. Alle hensynene er relevante i forhold til prosjektet, vurderingen vil bestå i å finne ut hvordan de skal prioriteres i forhold til hverandre. Man kan velge å la de rent språklige hensynene veie tyngst, da blir ordboken ytre sett mest systematisk og ryddig. Men en kan også velge å la brukeraspektene være styrende for valg av ekvivalent. Da vil transparens bli avgjørende. Et viktig hensyn er det at de ekvivalentene som velges, er i samsvar med usus i nordisk leksikografi. En ekvivalent som *lemmaoppløsning* er sannsynligvis mindre ekvivalent med utgangslemmaet enn de øvrige forslagene, men samtidig er den mer opplysende for norske brukere av ordboka, og den er mer i samsvar med faktisk brukt terminologi i norsk leksikografi. Man kan derfor spørre seg hva som er best, *delemmatisering*, som samsvarer best med utgangslemmaet, *avlemmatisering*, som vel er lettere forståelig, men som faktisk er et hybridord, noe vi etter redaksjonsprinsippene skal unngå, eller *lemmaoppløsning*, som er delekvivalent med utgangslemmaet, men som er mest i samsvar med faktisk bruk, i hvert fall i norsk leksikografi.

Det letteste er ofte å velge en ekvivalent som bare er en adaptasjon av utgangslemmaet

i kildespråket. Det vil si at det lages et nytt lånord, og dermed unngår man alle konnotasjoner og den upresisjon som måtte hefte ved en mer målspråksnær ekvivalent, samtidig som det er internasjonalt kjent og forstått. Et aktuelt eksempel på dette er nettopp termen *lemma*. I norsk tradisjonell leksikografi har en brukt *oppslagsord*, som er en partiell ekvivalent til *lemma*. Men *oppslagsord* er ikke bare en term, det er også et allmennord, og er derved mindre spesifisert enn termen *lemma*. Det som taler mot å velge en ekvivalent som *lemma*, er bl a at det ikke er noe kjent ord, og den som skal sette seg inn i fagområdet, må dermed lære seg en ny term. Et annet problem med *lemma* er at det er mer brukt og har fått en egen definisjon innen en bestemt leksikografisk teori, og at en kanskje trenger kjennskap til hele denne leksikografiske teorien og dens metode for å forstå i sin fulle dybde hva som skiller *lemma* fra *oppslagsord*. (Og enda verre er det dersom en term har fått forskjellig definisjon innen forskjellige teorier!) Et tredje problem er at *lemma* har en tradisjonell betydning (valgspråk, motto, overskrift, innholdsfortegnelse), noe som kan føre til misforståelser, men det velger jeg å holde utenom her. For en nærmere beskrivelse av termen *lemma*, se ellers Werner Wolskis artikkel i HSK 5.1: *Das Lemma und die verschiedenen Lemmatypen*.

I stedet for å velge ekvivalenten *oppslagsord* til kildespråkstermen *lemma* kan en velge *grunnform*. Den sier mer om innholdet i *lemma*, og gir en som er ukjent med fagområdet lettere assosiasjoner til hva som ligger i termen. Men *grunnform* kan også føre på villstrå, da det brukes i diakrone språkbeskrivelser der en snakker om avledninger og andre utviklinger med opphav i en *grunnform*. I NLO har vi foreløpig bestemt å regne internasjonalismen *lemma* som hovedlemma, og å føre opp både *oppslagsord* og *grunnform* med henvisning til *lemma*, der definisjonen av termen skal stå. Det blir gjort fordi NLO blant annet skal gi hjelp ved tilegnelse av internasjonal metaleksikografi. Alternativt kunne en tenke seg at også termene *oppslagsord* og *grunnform* fikk en kort definisjon med henvisning til hovedoppslaget *lemma*, der definisjonen er mer utfyllende. Det vil bli et spørsmål om plass og ressurser, og om hvilken stil NLO skal ha. Noe som er åpenbart, er at valg av ekvivalent fastlegger mye av stilen i ordboka, og også i hvilken grad en velger å binde seg til de forskjellige metaleksikografiske arbeidene som fins, eller om en velger å bygge på de tradisjonene som en har innen nordisk leksikografisk arbeid. Et siktemål med NLO er at den skal virke normerende/standardiserende, samtidig som den skal være mest mulig deskriptiv, og det er da viktig å vurdere de forskjellige alternativene nøye, slik at termene faktisk blir akseptert og brukt.

I alt oversetterarbeid er det et kjent problem at form i kildespråk smitter over på form i målspråk, uten at det ligger noe bevisst valg bak. Det samme kan selvfølgelig skje i NLO, og derfor bør en være på vakt mot å adaptere internasjonalismene uten å kartlegge grundig om det foreligger en nordisk ekvivalent, eventuelt delekvivalent. Det er lett å bli bundet av kildelemmaet, noe som har vist seg ved flere tilfeller der vi har foreslått forskjellige ekvivalenter til samme lemma, avhengig av om kildespråket har vært engelsk, fransk eller tysk. Vi får da dubletter:

| | |
|--------------------------|--------------------|
| dictionnaire de décodage | → avkodningsordbok |
| passive dictionary | → passiv ordbok |

Disse ekvivalentene må sammenføres til ett begrep, slik at vi ikke ender opp med noe en kunne kalle falske lemnaer.

I tillegg til termene fra HSK har vi tatt inn leksikografiske termer fra nordiske språk, ekserpert fra nordisk leksikografi og metaleksikografi. Disse termene kan være synonyme - helt eller delvis - for ekvivalentforslagene for de internasjonale termene. Det blir da viktig å velge den bokmålsekvivalenten som best svarer til den internasjonale termen, og som samtidig harmonerer best med dansk, svensk og dels islandsk og nynorsk. Vi må ta stilling til om den engelske, tyske eller franske formen er forskjellige uttrykk for samme begrep, eller om to forskjellige former i samme språk er to uttrykk for samme begrep.

3 Manglende og mangelfull ekvivalens

Ekvivalens kan beskrives som et graduelt forhold fra total mangel til total ekvivalens. Total mangel har med kulturelle forskjeller fra språksamfunn til språksamfunn å gjøre. Fra leksikografien kan forholdet eksemplifiseres med det norske uttrykket *klammeform*, som i norske ordbøker betyr en "tillatt sideform innenfor den offisielle normen", det vil si en form som kan benyttes av private rettssubjekter, men ikke av offentlig tilsatte eller offentlige organer. Det er derfor ikke tilstrekkelig i dansk eller svensk med ekvivalenten *variantform*, det må i så fall en ekstra forklaring til, da *variantform* er mye videre enn norsk *klammeform*.

Ofte er det rent tilfeldig at en ikke har ekvivalentformer, i leksikografi blant annet på grunn av den manglende teoriutviklingen på nordisk område. Det er derfor mange termer som mangler ekvivalent, men som er helt nødvendige i metaleksikografisk litteratur. Skal f eks det tyske *Wörterbuchgegenstand* ha ekvivalenten *ordboksgjenstand*, *-formål*, eller *-objekt*? Det er vanskelig å se at det her er noen naturlig norsk ekvivalent, og vi har heller ikke funnet belegg på noen av de foreslåtte norske ekvivalentene. Men termen er selvfølgelig nødvendig og bør etableres i NLO. Spørsmålet er da om en term skal konstrueres, eller om en skal nøye seg med perifrase.

Det kan også være den faglige tradisjon eller utvikling som gjør at det ikke fins noen ekvivalent. For eksempel har tysk og andre "verdensspråk" mange flere ordbokstyper enn norsk:

| | |
|----|---|
| ty | Schüler-, Lern-, Lehr-, Studienwörterbuch |
| no | stor ordbok, skoleordbok |

Det må her vurderes om det skal etableres nye termer, f eks *ungdomsskoleordbok*, *innlæringsordbok*, *opplæringsordbok* eller *studieordbok* før selve fenomenet fins. Det kan ellers være litt tilfeldig at ekvivalenter mangler, f eks har tysk *Grosswörterbuch* ingen norsk ekvivalent, men kunne hete *storordbok*. At termen ikke fins på norsk, skyldes nok mangel på slike ordbøker.

Ellers kan særpråklige tradisjoner være årsak til manglende ekvivalens. Det tyske *Grosschreibung*, engelske *capitalization*, har vi foreslått å kalle på norsk *bruk av stor bokstav*. Fenomenet fins selvsagt på norsk/nordisk, men det har ikke vært gjenstand for

særlig diskusjon, og dermed har man heller ikke hatt behov for en særskilt term. Det må da avgjøres om vi skal lage en, f.eks. *versalisering*, om termen bare skal forklares, eller om den ikke skal være med i det hele tatt.

Før akronymer er det ofte umulig å finne ekvivalenter, f.eks. *EFL - English for Foreign Learners*, eller *SGML - standard generalised mark-up language*. At det ikke er mulig å finne ekvivalenter her, viser kanskje først og fremst at de ikke bør være med i en ordbok som NLO, det er bare mulig å gi encyklopedisk informasjon om slike akronymer.

Dersom det for en internasjonalisme mangler en nordisk ekvivalent, er det foreløpig bestemt at det lages en adaptert internasjonal ekvivalent. Det er det eneste naturlige for de helt rene fagtermene, der det aldri har forekommet noen nordisk ekvivalent, slik som ved *glossering*. Men ofte vil det finnes en naturlig nordisk ekvivalent også, selv om den ikke er innarbeidet som term:

| | |
|----------------|------------------------------------|
| microstructure | → mikrostruktur / artikkelstruktur |
| macrostructure | → makrostruktur / ordboksstruktur |
| megastructure | → megastruktur / rammestruktur |

Det kan her diskuteres om det foreligger manglende ekvivalens, for de nordiske ekvivalentene er jo naturlige nordiske ord, men har vel ikke vært brukt som fagtermer. Det er ofte et problem å skille mellom fagspråk og allmennspråk, og for å markere skillet kan det være hensiktsmessig å velge bare internasjonalismene. Det er likevel ikke alltid gjennomført i de ekvivalentvalg som er foreslått:

| | |
|----|--------------------------------------|
| ty | Aussentext |
| fr | text externe |
| en | outside matter |
| bm | ekstern tekst / ytre tekst / omtekst |

Ifølge prinsippene skulle *ekstern tekst* være den beste ekvivalenten, men det kan hende at *ytre tekst* er vel så dekkende og forståelig, selv om det neppe fins belegg for uttrykket som term i nordisk metaleksikografi. Antakelig har det sammenheng med den tidligere nevnte manglende teoriutviklingen på nordisk grunn. Eksemplet viser for øvrig at det ikke alltid er lett å følge generelle prinsipper i leksikografisk arbeid. Der en nordisk term er innarbeidet, tas den selvfølgelig med, det er jo da heller ikke snakk om manglende ekvivalens. Det vil likevel være en avveining hvor godt kjent en nordisk term skal være for at den skal kunne regnes som en ekvivalent. Som mulige ekvivalenter til termen *encyclopaedic dictionary*' fins både *encyklopedisk ordbok* og *allkunneordbok*, men den sistnevnte formen er nok bare kjent som term innen visse miljøer i norsk leksikografi, og vil muligens gi akseptabilitetsproblemer.

4 Partiell ekvivalens

I terminologisk arbeid er det mulig å definere seg til fullstendig ekvivalens, siden defi-

nisjonene kan gjøres stipulativt. Men dersom NLO skal være deskriptiv, får vi mange tilfeller av partiell ekvivalens.

Delvis ekvivalens skyldes ofte perspektivforskjeller, som *ordboksartikkel* og *inngang* for tysk *Wörterbucheintrag*. Hvor mange perspektivmarkeringer en kan føre opp som egne lemmaer, bestemmer hvor utfyllende og omfangsrik ordboka skal bli. Dette har selv sagt med lemmautvalget å gjøre, men får konsekvenser for ekvivalensspørsmålet, da en ved å sammenføre lemmaer som uttrykker forskjellige perspektiver, bare oppnår en delvis ekvivalens i termene.

Et annet problem er det når en term har en "overflateekvivalent" i målspråket, men den er lite brukelig fordi den er svært etablert på andre bruksområder eller er beheftet med konnotasjoner. Dette forholdet drøfter Grindsted (1988) som det hun kaller pragmatiske ekvivalensrelasjoner. For eksempel har dansk satt opp ekvivalenten *anneks* for tysk *Annex*, mens svensk og norsk har *tillegg*. Den allmennspråklige betydningen vil nok virke forstyrrende inn på termen *anneks* i norsk. Men dersom vi velger *tillegg* for norsk og svensk, hva gjør vi da med *annexierte Mikrostruktur*? Å velge en ekvivalent for en term gir ofte ringvirkninger på andre.

5 Språkssystematiske forskjeller

Andre ekvivalensproblemer kan komme av de enkelte språkernes leksikalske struktur. Termer bør helst være substantiver, men noen språk danner substantiver lettere enn andre. Det er for eksempel ikke lett å finne et substantiv på norsk som dekker det tyske *Bidirektionalität*. Hvis vi ønsker en nordisk ekvivalent her, må innholdet omformes til et adjektiv, f.eks. *toveis*, og fenomenet kan da ikke føres opp og defineres isolert, bare som del av et nominalhypotagme, f.eks. *toveis ordbok*. Norsk kan ikke danne adjektivabstrakter som **toveishet*. Vi finner dermed ikke en naturlig ettordsterm for dette begrepet i norsk, hvis vi ikke velger internasjonalismen *bidireksjonalitet*.

Det er et ønske, om ikke nødvendigvis et ufravikelig krav, at termer består av ettordsenheter. Språk har ulik evne til å danne slike ettordsenheter, og det kan igjen gi ekvivalensproblemer. Det tyske *Ausgangssprachenseite* kan ikke omsettes til *utgangsspråksside*, som er den direkte oversettelsen. En må derfor velge termen *kildespråk*, men det er videre enn *Ausgangssprache*. Det tyske *sprachübliches Wort* må oversettes med *vanlig ord*, da **språkvanlig ord* ikke fins. Men *vanlig ord* er mindre spesifisert, da det ikke sier noe om forholdet til språket som system, slik *sprachübliches Wort* gjør.

Det kan være språkssystematiske forskjeller som gjør at ett språk har en mer finsiktet terminologi enn et annet. Det engelske *accent mark* bør på norsk ikke oversettes til *aksentmarkering*, men til *trykkmarkering*. *Aksentmarkering* er allerede i bruk for tonelagsmarkering, som også kalles aksent. Norsk må altså skille mellom aksent og trykk. Dermed spørs det om ikke også dansk og svensk bør la være å velge ekvivalenten *aksentmarkering* for *trykkmarkering*, ellers blir det uforklarlige og uheldige forskjeller mellom de nordiske språkene.

6 Kulturbundne forskjeller

Ekvivalensproblemer kan være uttrykk for kulturspesifikke forskjeller. Det gjelder selv sagt også for språkkulturelle forskjeller. En engelsk *Dictionary of Foreign Word* er noe annet enn det vi på norsk kaller *fremmedordbok*, og gir bare delvis ekvivalens. Engelsk er jo mer et blandingsspråk, mens norsk har et mindre sammensatt opphav, så *fremmedord* er mye videre enn *foreign words*. Likevel er det vel ikke mulig å finne en annen brukbar ekvivalent. Det gjenstår å bestemme hvor mye tilleggsinformasjon om slike forhold NLO skal gi.

7 Harmonisering mellom de nordiske språkene

I forbindelse med ekvivalentvalg dukker spørsmålet om harmonisering mellom de nordiske språkene opp. Et mål i terminologisk arbeid er økonomisering. Det vil bli å si at redundante ledd i formene ikke bør forekomme. Men av hensyn til konvensjoner i de enkelte nordiske språkene kan det være vanskelig. Ofte er det ledd som markerer et visst perspektiv, som kan betraktes som redundante:

| | |
|-------|---------------------|
| no,da | betydningsoppdeling |
| sv | betydelseindeling |

Kunne man her redusert til *betydningsdeling* for alle tre språkene, eller er det betydningsskillende om det står *opp* eller *inn* her?

Ellers kan det være minimale morfologiske forskjeller mellom språkene i de foreslåtte ekvivalentene. Oftest er det slik at svensk skiller seg ut fra norsk og dansk:

| | |
|-------|--------------------|
| no,da | alfabetisk ordning |
| sv | alfabetisk orden |

Strengt tatt er disse ikke ekvivalente. Semantisk er det vel her en forskyvning i perspektiv fra prosess til produkt, og dette perspektivskillet kan ikke være språkhengig, så det kan være at vi her har å gjøre med individuelle eller personlige tilbøyeligheter. Ofte forekommer forskjellige former som utganglemma, her *alphabetic ordering* og *alphabetic organization*, og det må klargjøres om dette er ett eller to lemmaer. Det viser at det er vanskelig å behandle ekvivalensproblemet uavhengig av lemmaseleksjon. Videre er det også umulig å fastlegge ekvivalentene før definisjonene er skrevet, noe bli a Picht & Draskau (1985) peker på. De sier at definisjonene er det eneste pålitelige instrument for å avgjøre ekvivalens. At to begreper likner på hverandre, er selvsagt ikke nok.

Et språk som har to variantformer, vil også ha interne ekvivalensproblemer. Det er egentlig minimale forskjeller mellom norsk bokmål og nynorsk, svært ofte fins det samformer. Men det å velge en samform vil samtidig være stilvelgende. På bokmål er formene *avgrenset ordbok* og *avgrensa ordbok* sidestilt, mens nynorsk bare kan ha *avgrensa ordbok*. Det samme gjelder for former som *rededannelse/reirdanning*.

Det å velge samform vil i mange tilfeller bety at en velger termer som ikke er i samsvar med språkbruk og språkfølelse hos mange. Dessuten vil samformene ofte skille seg fra

de danske, og til dels de svenske, men bokmålets særformer kan gi samform med dansk eller svensk. En må derfor foreta en avveining mellom fordelene ved en norsk samform som skiller seg fra de øvrige nordiske formene, eller velge en bokmålsform som skiller seg fra den nynorske, men som harmonerer med de øvrige nordiske. Dette er et språkpolitisk spørsmål som det ville være interessant å få synspunkter på, da det må løses på en eller annen måte, helst slik at ordboka blir mest mulig økonomisk redigert.

8 Avslutning

Et problem i all tospråklig leksikografi er hvordan en kan avgjøre hva som er ekvivalent. På dette innledende stadiet i arbeidet med NLO er det lett å bli veldig opptatt av formelle forhold, men det er selvfølgelig det å finne den mest dekkende ekvivalenten som er den viktigste og vanskeligste oppgaven. Og det er et vanskelig spørsmål å avgjøre hvem som er best kvalifisert til å vurdere det. I translatologi regner Koller den såkalte ideale tospråklige taler som den rette til å vurdere ekvivalens, uten at han går nærmere inn på hva det innebærer. I forbindelse med terminologi er selsagt situasjonen enklere i og med at definisjonene kan gis stipulativt. Men da NLO har som siktemål å være mest mulig deskriptiv, vil ekvivalensvalget være avhengig av god kunnskap om både kilde-språk og målspråk, samtidig som kunnskap om leksikografi, både teoretisk og praktisk, er en nødvendig forutsetning. Støtte og innspill fra hele det nordiske leksikografiske fagmiljøet er derfor ønskelig!

9 Litteratur

- Gasthuber, H. 1985 *Grundsätze, Ergebnisse und Tendenzen der Terminologienormung in Österreich*. I: Terminologie und Benachbarte Gebiete 1965-1985. Infoterm Wien
- Grindsted, A. 1988. *Prinsipper for præsentation af ækvivalenter i oversættelsesordbøger*. Odense
- Koller, W. 1992. *Einführung in die Übersetzungswissenschaft*. Berlin
- Picht, H. & Draskau, J: 1985. *Terminology: An Introduction*. København
- Wolski, W. 1983. *Das Lemma und die verschiedenen Lemmatypen*. I Hausmann & al: Wörterbücher... Handbücher zur Sprach- und Kommunikationswissenschaft Band 5.1. Berlin, New York

Ordbogsredaktion med COMPULEXIS hos Munksgaard

Anders Geertsen

Ethvert ordbogsprojekt starter med at forlag, hovedredaktør og forfatterne diskuterer redaktionsprincipperne for ordbogen. Disse redaktionsprincipper skrives ned og redigeres, så de mere og mere udgør en redaktionsvejledning, dvs et dokument, som kan blive en daglig manual for alle involverede i ordbogen, men mest af alt naturligvis en manual for forfatteren/forfatterne. En vigtig opgave for redaktionsvejledningen er at fastlægge principperne for hvordan de enkelte ordbogsartikler opbygges.

Feltopdelt data

Helt grundlæggende betragter vi en ordbogsartikel som en samling mindre elementer, hver med deres leksikografiske indhold. Hvis man arbejder med databaser ville man sige, at hver artikel er en "record", og at denne record kan bestå af en række "felter". En typisk ordbogsartikel kan fx se således ud i den trykte ordbog:

arm³ *vt & refl* 1. bevæbne, opruste 2. forsyne, udruste (*armed with the truth*) 3. armere, klargøre (*fx bombe*) ■ *in arms* under våben, (op)rustet; *lay down one's arms* nedlægge våbnene; *take up arms* gribe til våben; *under arms* under våben; *be up in arms* (sædv. fulgt af *against, about*) øve væbnet modstand imod

Når forfatterne og forlaget arbejder med denne artikel vil vi samtidig betragte den som en sekvens af felter, hver med deres specielle type information. Hvert felt har i følgende eksempel en lille tre-bogstavskode:

| | |
|-----|---------|
| HWD | arm |
| HOM | 3 |
| PSA | vt |
| PSA | refl |
| LV3 | 1 |
| TSL | bevæbne |
| TSL | opruste |
| LV3 | 2 |
| TSL | forsyne |

| | |
|-----|--------------------------------------|
| TSL | udruste |
| EXA | armed with the truth |
| LV3 | 3 |
| TSL | armere |
| TSL | klargøre |
| SEB | fx bombe |
| IDM | in arms |
| TSL | under våben |
| TSL | (op)rustet |
| IDM | lay down one's arms |
| TSL | nedlægge våbnene |
| IDM | take up arms |
| TSL | gribe til våben |
| IDM | under arms |
| TSL | under våben |
| IDM | be up in arms |
| GRA | sædv. fulgt af <i>against, about</i> |
| TSL | øve væbnet modstand imod |

En ordbogs redaktionsvejledning skal helt overordnet fastlægge og beskrive hvilke typer elementer eller felter, der skal kunne indgå i en typisk artikel. Visse felter vil være til stede i alle artikler, mens andre, fx feltet IDM med eventuelle idiomatiske udtryk, kun vil optræde i nogle artikler. Samtidig beskriver redaktionsvejledningen hvordan indholdet af de forskellige felter vil komme til at se ud på tryk i den færdige ordbog.

Mens det er en forholdsvis enkel opgave at fastlægge hvilke felter, der skal anvendes for at kunne opbygge de enkelte artikler i ordbogen, er det en langt mere kompliceret opgave at opstille et fornuftigt og entydigt regelsæt for rækkefølgen af felter, iterationer af felter, hierarkier og autoriseret input i de enkelte felter. Lad mig give eksempler på disse fire problemstillinger:

Rækkefølgen af felterne. Eksempel: Et idiomatisk udtryk (IDM) vil ofte være efterfulgt af en grammatisk oplysning (GRA), og derefter af en oversættelse (TSL). Vi har derfor en typisk sekvens af tre felter: IDM + GRA + TSL:

| | |
|-----|--------------------------------------|
| ... | |
| IDM | be up in arms |
| GRA | sædv. fulgt af <i>against, about</i> |
| TSL | øve væbnet modstand imod |
| ... | |

Hvis denne sekvens optræder hyppigt, og hvis vi bevidst ønsker netop denne rækkefølge, frem for en anden, må vi beskrive den i redaktionsvejledningen, og eventuelt sikre, at vores software kan genkende den og favorisere den frem for andre mønstre af felter.

Iterationer af mindre grupper af felter. Eksempel: En artikel vil i mange tilfælde blive opdelt i flere mindre afsnit, hver indledt af et arabertal. Man kan derfor betragte arabertallet med dets efterfølgende oversættelser som en fasttømret gruppe, som kan gentages efter behov:

...
 LV3 1
 TSL bevæbne
 TSL opruste

LV3 2
 TSL forsyne
 TSL udruste

...

Når man opbygger nye ordbogsartikler, eller redigerer i bestående, skal man være klar over, at et enkelt element i en sådan fasttømret gruppe ikke kan slettes, flyttes eller forandres, uden at det har indflydelse på de andre elementer i gruppen. Til gengæld kan en sådan gruppe ofte flyttes til et andet sted i artiklen, blot man flytter alle felter i gruppen som en samlet blok. Igen skal redaktionsvejledning og software gøre forfatteren opmærksom på dette.

Hierarkiet af felter. Eksempel: Visse felter, eller grupper af felter, er underordnet andre felter, eller grupper af felter. Et brugseksempel tilknyttet en arabertalsgruppe er underordnet denne gruppe. Denne relation mellem overordnede, sideordnede og underordnede elementer kan måske bedst illustreres ved hjælp af indryk:

HWD arm
 HOM 3
 PSA vt
 PSA refl

LV3 1
 TSL bevæbne
 TSL opruste

LV3 2
 TSL forsyne
 TSL udruste

EXA armed with the truth

...

Forholdet mellem over-, side- eller underordning af elementer er centralt for den typografiske

Anders Geertsen

grafiske manifestation af ordbogsartiklen. Populær sagt vil underordnede elementer ofte arve typografiske egenskaber fra overordnede elementer.

Autoriseret input i visse felter. Eksempel: Oplysninger om opslagsordets ordklasse hører hjemme i feltet "PSA" (Part of Speech A-language):

```
...
PSA      vt
PSA      refl
...
```

En opgave for redaktionsvejledning og software er naturligt nok at fastlægge hvilke forkortelser der anvendes for hvilke ordklasser, dvs fastlægge autoriseret input i PSA felterne.

Redaktionsvejledningen

Artiklernes opbygning, anvendte felter, typiske sekvenser og iterationer af felter, hierarkier m.v. forklares og beskrives grundigt og ved hjælp af mange eksempler i redaktionsvejledningen. Vi gør et stort arbejde for at skrive så grundige og detaljerede redaktionsvejledninger som muligt. Brugen af hvert enkelt felt forklares indgående, og vi giver eksempler på typiske sekvenser af felter, iterationer, mønstre m.v.. Det følgende viser et udsnit fra en redaktionsvejledning, som beskriver hvordan opslagsord, som er forkortelser, behandles ved hjælp af felterne FFA:

"Feltet Full Form anvendes når et opslagsord eller en oversættelse er en forkortelse. Både for opslagsord, som er forkortelser, og for oversættelser, som er forkortelser, kan det være ønskeligt at kunne angive den fulde form.

Til den fulde form af A-sprogs ord (dvs opslagsord) anvendes feltet FFA. Til den fulde form af B-sprogs ord (dvs oversættelser) anvendes feltet FFB:

I CX-Basic:

```
      HWD      EEC
      PSA      forkort
->FFA      European Economic Community
      TSL      EØF
->FFB      Det Europæiske Økonomiske Fællesskab
```

I den trykte ordbog: **EEC** *forkort = European Economic Community* **EØF** *= Det Europæiske Økonomiske Fællesskab"*

Ensartet feltstruktur for alle vores ordbøger

Selv om ordbøger er - og skal være - forskellige, har de mange iboende lighedspunkter. Enhver bilingval ordbog opererer med et A- og et B-sprog, med oversættelser (ækvivalenter), brugseksempler og idiomatiske udtryk. Ordklasseoplysninger og semantiske etiketter vil også være at finde i de fleste tosprogsordbøger. Vi kan derfor i meget høj grad ensrette brugen af felterne på tværs af alle vores ordbøger. Konkret betyder det, at et idiom altid står i et IDM felt, og at en ordklasseoplysning for A-sprogsord altid står i et PSA felt osv. Denne enighed gælder ikke blot feltnavnene, men også hele regelsættet for typiske iterationer, sekvenser af felter og hierarkiske relationer mellem felter. Og enigheden gør det muligt at lagre alle vores ordbøger i én fælles konsensus-database. Sammenligninger af lemma-udvalg eller idiomatiske udtryk, vending af ordbøger m.v. kan foretages ud fra denne konsensus-database.

Redaktionsværktøjer

Alle Munksgaards ordbøger bliver redigeret og lagret i et specielt ordbogssystem, COMPULEXIS Dictionary System. Dette system, der er udviklet af COMPULEXIS i Oxford, har to hovedelementer: **CX-Basic**, der bygger på WordPerfect 5.1 og som bruges til indtastning/redigering af data, og **CX-Master**, som bygger på en ORACLE SQL database, og som bruges til at samle alle ordbogsdata, vedligeholde konsensus-databasen, generere korrekturprint og levere data til sats.

CX-Basic

CX-Basic er indtastnings- og redigeringsdelen af COMPULEXIS. Forfatterne vil normalt arbejde i CX-Basic. Kernen i CX-Basic er tekstbehandlingsprogrammet WordPerfect, version 5.1. Ved hjælp af en række makrofunktioner er denne WordPerfect-udgave specielt velegnet til at redigere feltopdelte data. For en forfatter ser en ordbogsartikel således ud i CX-Basic:

| | |
|-----|----------------------|
| HWD | arm |
| HOM | 3 |
| PSA | vt |
| PSA | refl |
| LV3 | 1 |
| TSL | bevæbne |
| TSL | opruste |
| LV3 | 2 |
| TSL | forsyne |
| TSL | udruste |
| EXA | armed with the truth |

...

CX-Basic kan installeres og afvikles fra enhver IBM-kompatibel pc med DOS styresys-

Anders Geertsen

stem. CX-Basic gør det muligt dels at redigere i eksisterende ordbogsartikler, dels at indtaste nye ordbogsartikler. Da CX-Basic er baseret på WordPerfect 5.1 indeholder CX-Basic ikke nogen egentlig database. CX-Basic producerer almindelige WordPerfect tekstfiler med feltopdelt data, som er beregnet til at blive importeret i CX-Master databasen, som er installeret hos Munksgaard.

CX-Master

CX-Master er den del af COMPULEXIS, hvor alle ordbogsdata lagres, alfabetiseres og systematiseres, og hvorfra forlaget kan generere satsfiler til de trykte ordbøger. Samtidig byder CX-Master på en række søgemuligheder, som gør det let at rette fejl i data, checke krydsreferencer m.v.

I CX-Master gemmes alle ordbogsdata i en SQL database. Alle de ordbogsdata, som redigeres af forfatterne i CX-Basic, vil blive importeret i CX-Master, og her gemt i databasen. Samtidig kan CX-Master eksportere data til CX-Basic i form af WordPerfect filer.

CX-Master kan vise - og printe - hvordan de ordbogsfiler, som forlaget har modtaget fra forfatterne i CX-Basic format, vil se ud på tryk i den færdige ordbog. Denne funktion anvendes bl.a. til at generere korrekturprint til forfatterne i takt med at arbejdet skrider frem. En anden mulighed er at lade forfatterne selve generere disse korrekturudskrifter direkte fra CX-Basic.

Manifestationslogik

Både i CX-Basic og i CX-Master indtastes al tekst i ordinær, uden angivelse af fx fed, kursiv m.v. Tilsvarende skal forfatterne normalt ikke indsætte nogen form for tegnsætning: punktum, komma, semikolon, parenteser og lignende. Såvel skriftsnit som tegnsætning genereres automatisk af CX-Master eller CX-basic når filerne printes. Skriftsnit og tegnsætning skabes på basis af en analyse af:

- Feltkoden.
- Feltets placering blandt andre felter, dvs en analyse af iterationer, sekvens og hierarki.

Da hver leksikografisk oplysningstype har sit eget felt, kan man til enhver tid ændre skriftsnittet for en bestemt type oplysning. Eksempelvis sætte alle ordklasseangivelser i parentes - eller fjerne parentesen og sætte dem i petit.

Typisk arbejdsproces

Redaktionen af en ordbog består typisk af følgende etaper:

- a) Eksport af data fra CX-Master hos Munksgaard

Munksgaard opbygger og vedligeholder i CX-Master databaser med bl.a. ordbøgerne i serien Munksgaards Ordbøger. I mange tilfælde vil dele af disse data skulle indgå i revisionen af en ordbog. Eller som halvfabrikata i en ny ordbog. Ordbogsarbejdet vil derfor ofte starte med eksport af ordbogsdata fra CX-Master.

Data bliver eksporteret i det WordPerfect format, som direkte kan anvendes i CX-Basic. Samtidig bliver data delt op i en række mindre filer, typisk à 100 Kb, så behandling af filerne i CX-Basic hos forfatterne bliver hurtig og bekvem.

b) Import af data i CX-Basic

Forfatterne vil modtage ordbogsartiklerne som CX-Basic filer (WordPerfect format) på disketter fra Munksgaard. Med en simpel DOS kommando lægges disse filer ind på harddisken på forfatterens pc.

c) Redaktion af data i CX-Basic

Forfatterne redigerer de enkelte artikler i CX-Basic, på egen pc. CX-Basic indeholder redigeringsfaciliteter, som bl.a. giver mulighed for:

- Redaktion af eksisterende ordbogsartikler og redaktion af enkelte felter i artikler.
- Sletning af artikler eller enkelte felter.
- Oprettelse af nye felter og nye artikler.
- Blokfunktioner for kopiering og flytning af felter eller hierarkiske blokke af felter, inden for én artikel, eller mellem artikler.
- Automatisk check af korrekt input i visse felter.

Forfatterne behøver ikke sikre en korrekt alfabetisk orden i artiklerne i CX-Basic. Når CX-Basic filerne bliver importeret i CX-Master alfabetiseres alle artikler på ny. Tilsvarende skal forfatterne ikke på nogen måde angive tegnsætning eller skriftsnit i CX-Basic filen.

d) Korrektur fra CX-Master, hos Munksgaard

Så snart en forfatter har afsluttet redaktionen af en fil i CX-Basic, sendes denne fil på diskette til Munksgaard, hvor der udskrives et korrekturprint på papir. I andre tilfælde vil vi lade forfatteren selv generere dette korrekturprint direkte fra CX-Basic. På dette print vil de enkelte ordbogsartikler være opstillet med korrekt skriftsnit og tegnsætning. Feltkoderne vil ikke være at se.

e) Rettelse af fejl i CX-Basic

Når forfatteren har korrekturprintet i hånden vil han kunne indtaste eventuelle rettelser og tilføjelser direkte i CX-Basic.

f) Satsfiler til den trykte ordbog

Når alle ordbogsartikler er redigeret færdig genererer Munksgaard de endelige satsfiler ud fra CX-Master basen.

- Og så til de spændende spørgsmål

Når man diskuterer feltopdelte data, redaktionsværktøjer og stringent kodning af leksikografiske data dukker der ofte de samme spørgsmål op igen og igen. Jeg vil her til slut prøve at svare helt punktuelt på et par af de mest hyppige spørgsmål:

I hvor høj grad sikres konsistente data i input-ledet?

Data indtastes og redigeres i CX-Basic hos forfatteren. I udformningen af CX-Basic er der sket en afvejning mellem forskellige hensyn. Det ene hensyn er til konsistente data. Her ønsker vi naturligvis at artiklerne i så høj grad som muligt skal leve op til redaktionsvejledningens anbefalede mønstre af felter. Vi sikrer dette ved at indbygge en række automatiske check i CX-Basic. Disse check omfatter især:

- Check af om feltnavnet er korrekt.
- Check af indhold i felter, hvor kun et kontrolleret vokabular er accepteret.
- Check af om visse centrale felter står korrekt i sekvensen af felter.
- Check af nummereringen af bogstavsektioner, arabertals- og romertalsafsnit.
- Check af hierarkier af felter.
- Check af om grupper af felter, som naturligt hører sammen, flyttes samlet ved "cut & paste"-operationer.

Disse check er dog ikke så omfattende, at der er tale om en fuldstændig "live parsing" af artiklens struktur. Og her kommer vi til det andet hensyn. Det handler om at vi ønsker os enkle standardværktøjer, og høj grad af frihed til selv at kunne konfigurere disse værktøjer. Vi har udtrykkeligt ønsket at forfatterne skulle arbejde i et kendt stykke software, - og her er WordPerfect ideelt. Desuden kan det afvikles på selv meget billige - og lidt ældre - 286-maskiner. Skulle CX-Basic foretage en komplet live parsing ville det stille væsentligt større krav til maskinen, og det ville i det hele taget ikke kunne være baseret på et stykke standardsoftware som fx WordPerfect.

Endelig er der et andet aspekt af sagen: Det er vores erfaring at man får det bedste resultat ved at gøre en stor indsats for at skrive en god redaktionsvejledning, og ved at få forfatterne til at følge den, samtidig med at man giver forfatterne et værktøj med en høj grad af frihed og fleksibilitet. Resultatet er, at vi nu og da får artikler fra CX-Basic, som ved import til CX-Master-basen viser sig ikke at overholde den struktur, som redaktionsvejledningen foreskriver. Disse afvigelser bliver da enten rettet under selve importen til CX-Master, eller senere i korrekturfasen, idet en forkert kombination af felter vil få manifestationslogik-modulet i CX-Master til at generere en forkert grafisk opsætning af artiklen.

Kan forfatterne se den færdige artikel på skærmen?

Nej, forfatterne kan ikke på skærmen se hvordan en artikel vil se ud i den trykte bog. I CX-Basic ses alle artikler som et WordPerfect-dokument, bestående af sekvenser af felter. Til gengæld kan vi udstyre forfatterne med en print-facilitet, som gør det muligt for dem at printe de færdige ordbogsartikler direkte fra CX-Basic.

Dette er et punkt, hvor vi er noget afventende: På den ene side ønsker de fleste nye forfattere at få lov til at se den færdige artikel på skærmen. Fra andre sammenhænge er de jo vant til WYSIWYG-faciliteter. Men på den anden side ved vi erfaringsmæssigt at vi får de bedste ordbogsartikler ved at tvinge forfatterne til at tænke i elementer med leksikografisk indhold fremfor i typografisk fremtoning.

Er det SGML?

Både ja og nej. Den feltopdelte og generiske kodning af alle typer information i en ordbogsartikel er i princippet en typisk SGML-kodning. Men i dag vil man intet sted i CX-Basic eller CX-Master finde en egentlig DTD - Document Type Definition - som følger syntaksen fra ISO-standarden. Ikke desto mindre findes der noget, som ganske modsvarer en DTD. Først og fremmest den meget detaljerede redaktionsvejledning, som beskriver anvendelsen af hvert enkelt felt, samt typiske sekvenser, hierarkier og iterationer af felter. Dernæst indeholder CX-Master et modul, som omformer sekvensen af felter i enhver artikel til et grafisk skriftbillede. Dette kan kun lade sig gøre, fordi CX-Master indeholder en formaliseret beskrivelse af alle felterne og deres indbyrdes relationer. Denne beskrivelse svarer ganske til en DTD, også selv om den ikke er formuleret i den af ISO vedtagne DTD-syntaks.

Mot en mångdimensionell tvåspråkig ordbok

Maria Toporowska Gronostaj

1. Inledning

Den tvåspråkiga lexikologins främsta syfte är att etablera relationer och länkar mellan två lexikologiska system samt att dokumentera och exponera dessa på ett systematiskt sätt. Det är dessa relationer som ligger till grund för de fullständiga eller partiella ekvivalensrelationer som återges i tvåspråkiga ordböcker. Den bild som förmedlas i tvåspråkiga ordböcker är dock väldigt fragmentarisk och den primära länken, dvs betydelse-relationen, markeras vanligen genom angivelse av en eller flera ekvivalenter. (Med termen 'ekvivalent' avses här både fullständiga och partiella ekvivalenter om inte annat anges). Denna praxis medför att relevant lexikalisk information förblir implicit och oåtkomlig för ordboksanvändaren. För att avhjälpa problemet görs i detta förstudium¹ ett försök att explicit återge all lexikalisk information som kan leda till en bättre förståelse av målspråket och ge insyn i språkliga intuitioner som infödda talare besitter. Att integrera lexikologiska tankegångar med lexikografisk applikation förefaller vara ett naturligt sätt att angripa problemet. Detta förväntas resultera i en mera holistisk och explicit framställning av lexikalisk information i en tvåspråkig ordbok.

Denna strävan att förbättra och förfina ordbokens innehåll medför att dess volym ökar markant. För att göra innehållet lätthanterligt definieras noga dess makro- och mikrostruktur. I mikrostrukturen skiljs 13 moment åt. Flera av dessa moment tillför ny information i jämförelse med traditionella tvåspråkiga ordböcker. Dit hör definitioner, semantiskt och syntaktiskt orienterade kommentarer, valenser och information om semantiska fält. På det morfologiska planet anges ekvivalentens böjning samt dess semantiskt relevanta sammansättningar och avledningar. Ekvivalentens transkriberade och inspelade uttal, många exempel samt idiom med förklaringar tillför ett ytterligare stöd för praktiska tillämpningar. För att göra den stora mängden av information lättåtkomlig för användaren provas lagring av det lexikaliska innehållet i dataformat för att möjliggöra en selektiv informationssökning. Ordboksmodellen föreligger som HyperCard-applikation.

I fokus för beskrivningen står ett urval (ca 160) svenska och polska verb. Ett sådant kontrastivt studium av verb, genomfört enligt samma principer och med samma analysdjup i de båda språken, medför att skillnader av semantisk, syntaktisk och lexikalisk natur kan kartläggas systematiskt, vilket är till gagn för framställning av tvåspråkiga produktiva ordböcker för dessa språk.

2. Ordbokens struktur och innehåll

Datatekniken har öppnat nya möjligheter för utveckling av ordböcker. Lexikografen förfogar numera över ett medium som ger honom frihet att arrangera materialet på bästa möjliga sätt utan att tänka på volymbegränsningar. Denna frihet är viktig för införlivandet av lexikologisk kunskap i ordböcker, vilket är till gagn för ordboksanvändaren. Dessutom kan lexikografen aktivt vägleda ordboksanvändaren i ordsökningen så att den kan göras effektiv och träffsäker. Högst prioriteras här dock sådan information som bidrar till bättre förståelse av målspråket. Att omvandla en sökningsprocedur till en upptäcktsprocedur bör vara det främsta målet med tvåspråkiga ordböcker av den här typen. För denna ordboksmodell är detta mål vägledande och inverkar på utformningen av den aktuella ordbokens struktur och innehåll.

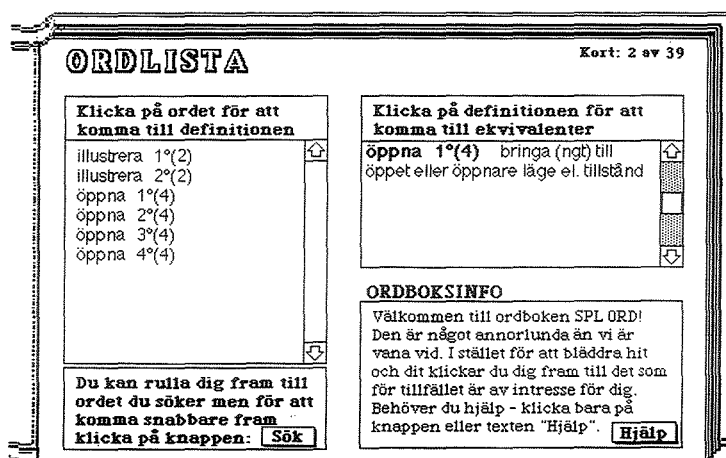
I och med att teoretiska och pragmatiska ändamål här flätats samman förefaller det vara motiverat att presentera ordboksmodellen i den form i vilken den vänder sig till sina användare, dvs användarens perspektiv väljs här som utgångspunkt för beskrivningen. Målgruppen är personer med svenska som modersmål som söker stöd vid genereringen av polska texter.

2.1. Ordbokens makrostruktur

Med hjälp av HyperCard som applikationsformat och HyperTalk som programmeringsspråk har en ordboksmodell konstruerats vars utformning påminner om en bunt kort. Varje korts innehåll och struktur är strikt anpassad till det aktuella sökningssteget.

(i) I det första steget presenteras ett titelkort som bekräftar den valda applikationen, dvs svensk-polsk ordbok: verb.

(ii) Från det första kortet går man automatiskt till nästa kort med beteckning ORDLISTA (Fig. 1.) som tillhandahåller en förteckning över de verb som finns med i ordboken.



Figur 1

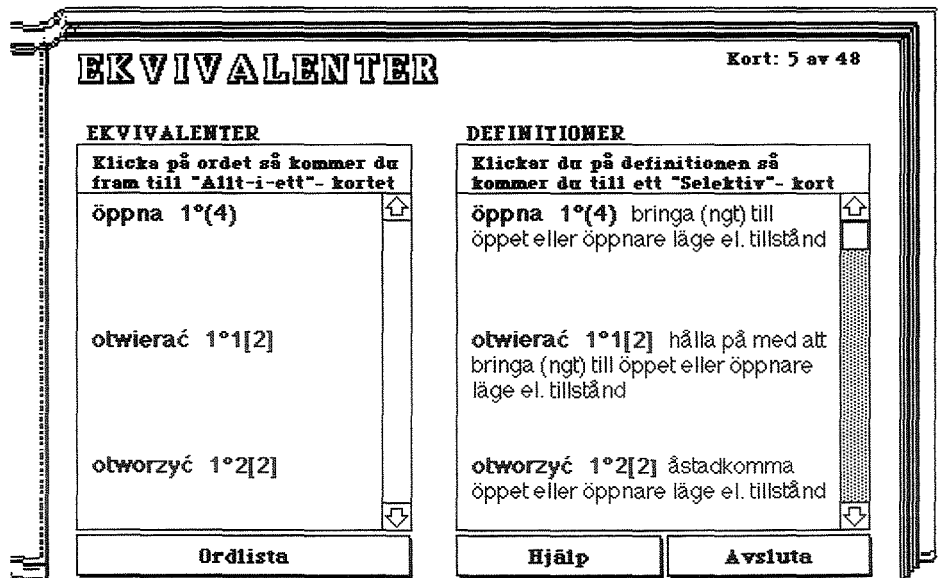
För att hitta det sökta verbet kan man gå till väga på två sätt. Man kan bläddra i det rullande fältet på kortets vänstersida eller anropa sökfunktion som letar efter lemmat² i fråga. Till höger om verbet finns en siffra som indikerar det löpande lexemnumret, följd av en siffra inom parentes som anger det totala antalet lexem³ upptagna i ordboken för verbet i fråga. (För en detaljerad beskrivning av lemma-lexem-modellen se Allén, 1971, Järborg, 1989.)

Det kan noteras att ekvivalenten har tre siffror: den första för att verifiera uppslagsordets och ekvivalentens lexemnummer, den andra för att indikera ekvivalentens löpande lexemnummer och den tredje för att ange det totala antalet lexem.

Genom att klicka på det utvalda verbet får man fram lemmats definition i fältet till höger. Proceduren kan uppreppas tills man hittat det sökta lexemet. Syftet med en sådan uppläggning av informationen är att redan vid den första kontakten precisera det lexem som söks och att därmed i de kommande stegen sortera bort den information som inte är relevant för den valda betydelsen. Genom att explicit formulera verbets definition och 'utsätta' ordboksanvändaren för dess innehåll aktiveras de betydelsefaktorer som gäller för uppslagsordet och som skapar en referensram för jämförelsen med de betydelsefaktorer som expliceras i ekvivalentens definition. (De svenska verbens definitioner är ofta identiska med definitionerna i *Svensk ordbok*.)

När man har hittat det sökta lexemet, klickar man in i definitionstexten för att få fram verbets ekvivalent(er).

(iii) Härmed aktiveras det tredje steget i sökningen som visas på det kort som kallas EKVIVALENTER (Fig. 2.) och som upprepar det svenska verbet tillsammans med dess definition samt lägger till ekvivalenter och deras definitioner.



Figur 2

Genom att samla all denna information på ett och samma kort får man facit till likheter och skillnader i betydelse. Utifrån denna information kan man välja den ekvivalent som motsvarar eller ligger närmast den sökta betydelsen. Utformningen av detta steg, liksom av det föregående, har dikterats av samma stora krav på sökningsprecision. Nu är de förberedande stegen avklarade och man kommer fram till den egentliga informationskällan om den utvalda ekvivalenten. Genom att klicka på ekvivalenten kommer man till ett ALLT-I-ETT-KORT (Fig.3.).

| | | | | |
|--|--|-----------------|--|--------------|
| öppna 1°(4) | | otwierać 1°1[2] | | PL |
| Definition | | | | |
| otwierać 1°1[2] hålla på med att bringa (ngt) till öppet eller öppnare läge el. tillstånd | | | | |
| Kommentar | | | | |
| Verb: kausativt Aspekt: imperfektiv Aktionsart: (1) Verb med fokus på ett pågående förlopp samt en antydning | | | | |
| Besläktade begrepp | | | | |
| töppna 1°(4) otwierać 1°1[2] = stänga zamykać | | | | |
| ↓ <u>glänta</u> (på dörren) uchylać (drzwi) | | | | |
| Besläktade ord | | | | |
| subst. öppnande otwieranie, (otworzenie); öppning otwarcie, (otworzenie), otwór, otworek; öppnare otwieracz; | | | | |
| Ordklass | | Böjning | | Uttal |
| Konstruktion Kort: 31 av 48 | | | | |
| Konstr. 1. S öppnade O (PO:med) (PO: åt / för) ... S öppnade (S-nom) otwierał (O-dat) O-ack | | | | |
| Exempel | | | | |
| 3. När mannen höll på att öppna dörren inträffade en explosion. -- Kiedy mężczyzna otwierał drzwi nastąpiła eksplozja. | | | | |
| Idiom och fraser | | | | |
| ~ dörren el. porten (för ngn / ngt) (bildl.) skapa möjligheter el. skydd för ngn el. ngt otworzyć komuś swoje podwoje; stworzyć komuś / czemuś | | | | |
| Ordlista | | Selektiv | | Hjälp |
| Ekvivalenter | | Avsluta | | |

Figur 3

Klickning på ekvivalentens definition gör att man hamnar på ett SELEKTIVKORT (Fig.4.).

| | | | | |
|--|--|---------------------|--|-----------------------|
| öppna 1°(4) | | otwierać 1°1[2] | | PL |
| Definition | | | | |
| otwierać 1°1[2] hålla på med att bringa (ngt) till öppet eller öppnare läge el. tillstånd | | | | |
| Kommentar | | Konstruktion | | Besläktade ord |
| Besläktade begrepp | | | | |
| töppna 1°(4) otwierać 1°1[2] = stänga zamykać | | | | |
| ↓ <u>glänta</u> (på dörren) uchylać (drzwi) = stänga till, dra till, tillsluta przymykać, domykać ↓ spärra upp (gapet, ögonen) otwierać szeroko (usta, oczy) ↓ gapa otwierać (szeroko) usta ↓ läsa upp otwierać kluczem, | | | | |
| Ordlista | | Ekvivalenter | | |
| Allt-i-ett | | Avsluta | | |

Figur 4

Skillnaden mellan korten ligger framförallt i layouten. I ett ALLT-I-ETT-KORT får man tillgång till all information samtidigt, med undantag av information om ordklass, böjning och uttal som finns gömd i sina lådor tills man öppnar dem genom klickning. På det selektiva kortet lämnas åt ordboksanvändaren att själv välja bland uppgifter rubricerade som kommentarer, konstruktion, besläktade ord, besläktade begrepp, exempel, idiom och fraser genom att klicka på den utvalda rubriken. Två fält kan visas åt gången och deras skrivvyta är större jämfört med den andra korttypen.

Som en extra informationskälla infogas i den tvåspråkiga ordboken en enspråkig ordbok för svenskan där det svenska verbets semantiska och morfosyntaktiska dimensioner presenteras. Dessa kort har samma layout som de ovan nämnda korten för ekvivalenter. Den information som är samlad på dessa kort ger en referensram för kontrastiva studier och gynnar därmed inte bara lexikografen utan också en normal ordboksanvändare.

2.2 Ordbokens mikrostruktur

Som redan har antytts i den operationella beskrivningen av makrostrukturen är ordbokens mikrostruktur explicit specificerad med rubriker som hänvisar till de dimensioner som aktiverats för verbet. Följande moment konstituerar en artikel:

- (1) Uppslagsord: uppslagsordet anges i grundform, dvs infinitiv för verb. Ord som är homografer markeras med upphöjd siffra före den första bokstaven.
- (2) Ekvivalent.
- (3) Ekvivalentens ordklass.
- (4) Ekvivalentens böjning.
- (5) Ekvivalentens uttal: transkription och inspelat uttal.
- (6) Ekvivalentens definition formulerad i källspråket.
- (7) Kommentarer av syntaktisk eller semantisk art förmedlade på källspråket. Bland de submoment som upptas här finns: aspekt, aktionsart, konstruktionsväxlingar, ellips etc.
- (8) Konstruktionsuppgifter: här anges semantisk och syntaktisk valens. På det syntaktiska planet anges dels den fullständiga valensramen med både obligatoriska och fakultativa fyllnadsled ifyllda, dels också den minimala ramen. Kontrastiviteten framgår tydligt i och med att valensramarna anges för källspråket och målspråket.
- (9) Besläktade begrepp aktiverar det semantiska fältet genom att ange hyperonymer, hyponymer, kohyponymer, synonymer och antonymer.
- (10) Besläktade ord dvs semantiskt relevanta avledningar och sammansättningar.
- (11) Exempel.
- (12) Idiom och fraser.
- (13) Språkkodning: PL eller S för polska respektive svenska.

Uppgifterna (8-12) har en tydlig kontrastiv prägel genom att hänvisa till både käll- och målspråket.

Med tillgång till alla dessa fakta bör man vara väl rustad för att korrekt generera texter i målspråket. För de användare som anser att en del fakta är överflödiga p g a att de redan besitter denna kunskap återstår alltid möjligheten att välja bort dem.

3. Teoretiska ramar för projektet

Frågan hur man kan förbättra ordbokens prestanda hänger samman med det teoretiska underlaget. Nedan skisseras några av de viktigaste frågeställningarna som ligger till grund för en lexikologisk analys i denna ordbok.

3.1 Betydelseanalys

Låt oss inleda detta avsnitt med Bo Svenséns (1987:134) syn på ekvivalens:

“Den tvåspråkiga ordbokens uppgift är att förse ord och uttryck på källspråket med semantiskt och bruklighetsmässigt så likvärdiga motsvarigheter som möjligt på målspråket. Men fullständig sådan överensstämmelse (ekvivalens) mellan ord och uttryck i två språk är ganska sällsynt, på samma sätt som det är sällsynt med fullständiga synonymer inom ett och samma språk.”

Av detta följer att betydelseanalys och valensuppgifter är av primär relevans för bestämning av ekvivalenter. I detta projekt bedrivs betydelseanalys med utgångspunkt i en beskrivning av verbets referentiella och relationella betydelse⁴. För att en fullständig ekvivalens ska råda måste verbens referentiella och relationella betydelser vara identiska. I annat fall har man att göra med en partiell ekvivalens.

Nu uppstår frågan om ordboksanvändaren utifrån informationen i ordboken kan fatta beslut om vilken typ av ekvivalensrelation som råder mellan uppslagsordet och ekvivalenten. Denna fråga kan besvaras med ja, eftersom den referentiella betydelsen kan spåras via definitionerna och den relationella finns angiven i valensuppgifterna.

Tack vare de stringenta regler som ligger till grund för definiering av ord i den vid Göteborgs universitet (Språkdata) upprättade lexikaliska databasen (LDB) får ordboksanvändaren insyn i ordbetydelsens lexikaliska dekomposition och de lexikaliska komponenternas plats i den interna hierarkiska betydelsestrukturen. (Se Järborg (1988) för information om definitionstagning enligt betydelsefaktorprincipen och Järborg (1989) om betydelseanalys och betydelsebeskrivning i LDB.) I och med att fullständig synonymitet och ekvivalens sällan förekommer förefaller det vara mycket angeläget att fånga de betydelsefaktorer och de betydelserelationer som skiljer käll- och målspråket åt.

Det finns även andra fördelar med att upptäcka betydelskillnader i början av sökningsproceduren därför att kunskap om dessa skillnader kan underlätta förståelsen av andra uppgifter i ordbokens mikrostruktur.

Till exempel föranleder betydelsefaktorn aspekt att antalet ekvivalenter ofta blir dubblerat jämfört med det svenska uppslagsordet. Redan detta faktum signalerar en partiell ekvivalens, vilket framhävs i definitionerna. Dessutom kan ekvivalenternas

benägenhet att kombineras med andra betydelsefaktorer variera. Detta gäller bl a aktionsart och aspekt. Dessa fakta är ofta svåra att hitta i grammatikor och tvåspråkiga ordböcker, vilket talar för att de bör tas upp i kommentarfältet. Verbets aspektuella karakteristika är inte mindre viktig för subklassifikationer inom verbets semantiska fält där perfektiva och imperfektiva verbformer skapar sina egna delmängder. Den aspektuella karakteristiken bör även beaktas på det morfosemantiska planet, där även deverbalerade substantiv kan signalera aspekt. Aspektuella distinktioner kan göra sig gällande på valensplanet när det gäller fyllnadsledens morfosyntaktiska form (ackusativ vs partitiv). De upptas i fältet med rubriken konstruktion. Dessa fakta talar för sig själva, vilket i det här fallet innebär att definitionen är ett legitimt moment även i den tvåspråkiga ordbokens mikrostruktur. Den bidrar till att etablera semantiska länkar mellan olika dimensioner och fångar därmed en mer holistisk bild av verbets betydelse och dess plats i det lexikologiska systemets olika dimensioner.

3.2. *Semantisk valens*

Av intresse för valensbeskrivning är främst de aspekter av referentiell och relationell betydelse som aktualiseras i samband med fastställandet av argument och deras selektionsrestriktioner. Med ett verbs semantiska valens förstås här verbets benägenhet att semantiskt implicera existensen av vissa typer av argument och därmed etablera de semantiska funktioner som råder mellan verbet och dess argument. Någon färdig uppsättning semantiska funktioner finns ej fastställd i detta skede av projektet. Däremot finns det intressanta förslag till sådana uppsättningar (Somers (1987), Allerton (1982), Brodda (1973)) och dessa kommer att utvärderas under projektets gång. Denna information kommer att kompletteras med uppgifter om argumentens selektionsrestriktioner.

Nästa analyssteg avser argumentens placering i en semantisk valensskala. Med hjälp av denna skala beskrivs verbets benägenhet att kvalitativt och kvantitativt presupponera sina argument, vilket leder till subklassificering och rangordning av argumenten. Den semantiska valensskalan består av följande steg:

- (i) *inherent argument*, som är högst rankade i skalan eftersom de komprimerar aspekter av den referentiella och relationella betydelsen (Text 'bil' är ett inherent argument för verbet *bila*.)
- (ii) *obligatoriska argument*, som står för nödvändiga medspelare i den av verbet betecknade situationen eller händelsen och som har status av en variabel.
- (iii) *fakultativa argument*, som varken är obligatoriska eller inherent men som är semantiskt förenliga med verbets betydelse.
- (iv) *alternativa argument* som i kontrast till de ovan nämnda typerna av argument uppvisar svaga semantiska band med verbet, om de jämförs med verbets övriga argument. Deras kännetecknande drag är att de potentiellt kan semantiskt alternera med

verbets inherenta eller obligatoriska argument. (Jfr verbet *bo* som egentligen kräver ett lokativt argument vilket dock kan alternera med ett argument som betecknar sätt.)

Även om ordboksanvändaren skonas från de metateoretiska övervägandena ovan så finns de dels implicit dels explicit i de uppgifter som förmedlas via ordboken. De manifesteras som skillnader i lexikaliseringen. Som exempel kan nämnas olika realisationer av det svenska inherent argumentet hos polska verb. För *bila* - *jechać / jeździć samochodem* 'åka bil' och *gapa* - *otwierać / otworzyć usta* 'öppna munnen' blir de externa argumenten obligatoriska. För verb som *klippa*, *skära* kan deras inherent argument omvandlas till obligatoriska om instrumentet står i fokus, annars förblir de fakultativa.

Den här kontrastiva aspekten framgår tydligt då parallella uppgifter för svenska och polska anges eller kommenteras. (För en mera utförlig analys av semantisk och syntaktisk valens se Toporowska Gronostaj (1991)).

3.3. Syntaktisk valens

Med syntaktisk valens menas de syntaktiska relationer som råder mellan verbet och dess fyllnadsled⁵. Fyllnadsledens funktion återges i en valensram med hjälp av sådana kategorier som valenssubjekt (S), valensobjekt (O), indirekt objekt (IO), reflexivt objekt (RO), prepositionobjekt (PO) för att nämna några. Vidare relateras dessa kategorier till Somers (1987) valensskala, som består av sex nivåer av vilka de första fyra (integrerade fyllnadsled, obligatoriska fyllnadsled, fakultativa fyllnadsled och mellanled) anses vara relevanta att tas upp i ordbokens valensramar.

Till integrerade fyllnadsled (integral complements) hör lexikaliserade uttryck i obestämd form, utan artikel, som inte tillåter strykning eller utbyte mot en synonym. De kan ej pronominaliseras och deras möjlighet att ta bestämning är starkt begränsad.

Med obligatoriska fyllnadsled (obligatory complements) åsyftar Somers i princip de led som semantiskt och morfosyntaktiskt determineras av verbet men som kan utelämnas vid nominalisering, passivisering eller omformulering till infinitiva konstruktioner.

De fakultativa fyllnadsleden (optional complements) uppfyller dels de krav som ställs på obligatoriska fyllnadsled, dels visar de benägenhet till utelämnning om pragmatiska skäl föreligger.

S k mellanled (middles) har gemensamma karakteristiska drag med både fakultativa fyllnadsled och de återstående två nivåerna i valensskalan, nämligen tilläggsled (adjuncts) och perifera tilläggsled (extra-periferals).

Som exempel på hur valensuppgifter återges i ordboken citeras här (Fig. 5.) de uppgifter som ges för verbet *öppna* 1°(4) - *otwierać* 1° 1[2].

Konstr. 1.

S öppnade O (PO:med) (PO: åt / för)

... S öppnade

(S-nom) *otwieral* (O-dat) O-ack (O-instr.)

... *otwieral*

Konstr. 2.

S öppnade O: refl. (PO: för)

a. (S-nom) *otwieral* O:się (P:przed + O-instr)

... (S-nom) *otwieral* O:się

b. (S-nom) O:się *otwieral* (P: przed + O-instr.)

... (S-nom) O:się *otwieral*

—
S-nom i (1) - /Agent/: varelse, person-ekvivalent, kraft, rörligt föremål, äv. abstrakt subst.

S-nom & O:się i (2) - /Agent/ el. Theme/: varelse, personekvivalent, (rörligt) föremål som kännetecknas av tillslutet läge el. tillstånd, äv. abstrakt substantiv

O-ack - /Theme/: spärrande anordning (dörr, lås...), föremål med tillsluten form, äv. abstrakt subst.

O-dat - /Beneficiary/: person, rörligt föremål

O-instr - /Instrument/: instrument, verktyg

P:przed + O-instr - /Beneficiary/: person, föremål

Figur 5

Med hjälp av dessa valenser förmedlas information om: (i) semantisk valens, (ii) syntaktisk valens och dess morfosyntaktiska exponenter, (iii) argumentens respektive fyllnadsledens obligatoriska (fet stil) eller fakultativa närvaro i valensramen, (iv) den maximala och den minimala valensramen, dvs argumentens benägenhet till ellips, (v) kongruensförhållande mellan fyllnadsleden (understrykning).

3.4. Information om verbets semantiska fält

Som redan har antytts utgör paradigmatiske relationer ett viktigt komplement till betydelseanalysen i och med att verbet presenteras utifrån ett bredare perspektiv, dvs som ett element i ett semantiskt fält. Med hjälp av sådana parametrar som synonym, hyperonym, hyponym, kohyponym och antonym definieras verbets plats i förhållande till övriga enheter i den semantiska referensramen. I en tvåspråkig ordbok aktiveras två semantiska system som i sin tur matchas mot varandra, vilket fångar och återspeglar skillnader och därmed det kontrastiva perspektivet. Det är det kontrastiva perspektivet som utgör den tvåspråkiga ordbokens kärna och som lyfter och vidgar dess innehåll.

I den här ordboken hittar man denna information under rubriken: besläktade begrepp. För verbet *otwierać* 1^o(4) vars definition lyder 'bringa (ngt) till öppet eller öppnare läge el. tillstånd' och dess imperfektiva ekvivalent *otwierac* 'hålla på med att bringa (ngt) till öppet eller öppnare läge' anges följande:

(↑ hyperonym, ↓ hyponym, = synonym, ≠ antonym)

Öppna 1°(4) otwierać 1°1[2]

≠ stänga zamykać

↓ **glänta** (på dörren) uchylać (drzwi)

≠ stänga till, dra till, tillsluta przymykać, domykać

↓ **spärra upp** (gapet, ögonen) otwierać szeroko (usta, oczy)

↓ **gapa** otwierać (szeroko) usta

↓ **låsa upp** otwierać kluczem, (otwierać z klucza)

≠ låsa zamykać na klucz

↓ **regla upp** odryglowywać

≠ regla zaryglowywać

↓ **knäppa upp** (knäppa 6°(6))rozpinać guzik(i)

≠ knäppa zapinać na guzik(i)

↓ **korka upp** odkorkowywać

≠ tillsluta, korka igen el. till zakorkowywać

↓ **bryta upp** rozrywać

≠ försluta zapieczętowywać, zamykać

↓ **sprätta upp** rozcinać

↓ **skära upp** rozcinać

↓ **dyrka upp** otwierać wytrychem

↓ **forcera** forsować

↓ **fälla upp** rozpinać

≠fälla ihop el. samman składać

↓ **spänna upp** napinać

≠ slaka poluzniać

Figur 6

Verben ovan har grupperats med hänsyn till realisationen av faktorn 'resultat' på följande sätt: (i) mer eller mindre öppet läge (ofta med avseende på spärrande anordning), (ii) öppet tillstånd med avseende på en lexikaliskt determinerad anordning (lås, knapp etc.), (iii) öppet tillstånd uppnått med hjälp av verktyg och/eller med inslag av våld, (iv) öppet tillstånd som förutsättning för funktionsduglighet e. d.

Det gäller generellt här att vid struktureringen av verbets semantiska fält beaktas följande: a) verbets referentiella betydelse, b) verbets relationella betydelse, c) verbets aktionsart samt dess implikationer för val av ekvivalenter och aspektuell form, d) verbets inherenta argument, e) verbets övriga betydelsefaktorer och deras fokusering.

4. Ordbokens lexikologiska profil vs lexikologisk tradition

Modellen som förespråkas här avviker avsevärt från den gängse tvåspråkiga lexikografiska tradition som i första hand strävar efter att förse ordboksanvändaren med lämpliga översättningsekvivalenter. Zgusta formulerar dessa tankegångar på följande sätt:

"The dictionary should offer not explanatory paraphrases or definitions, but real lexical units of the target language which, when inserted into the context, produce a smooth

translation. This is a perfectly natural requirement. Lexicographers have followed it since time immemorial." Zgusta (1984:147).

Zgusta (1987:31) är väl medveten om att ett sådant förhållningssätt medför begränsningar eftersom ordböcker konstruerade enligt denna princip bara kan användas för att översätta texter till modersmålet eller till målspråket under förutsättning att man redan behärskar målspråket i fråga.

Modellen som förespråkas här vänder sig i första hand till dem som håller på att lära sig målspråket eller vill fördjupa sin kunskap i målspråket. Den har många drag gemensamma med sk inlärningslexikon av typ LDOCE, COBUILD och är tänkt att aktivt stödja inläring av det främmande språket. För att göra den proceduren så friktionsfri som möjligt används genomgående källspråket som metaspråk. Denna typ av ordbok bör betraktas som ett komplement eller som ett alternativ till redan existerande tvåspråkiga ordböcker. På grund av den extensiva presentationen av verbets olika dimensioner lämpar den sig inte för tryck i pappersversion, men det finns möjlighet att skriva ut allt som står under respektive verb.

Denna extensiva presentationsform för innehållet kan ses som ett uttryck för att tvåspråkig lexikografi ej bör presupponera existensen av på förhand givna översättnings-ekvivalenter, utan snarare bör utgå från det motsatta och mera skeptiska antagandet att det mellan två lexikologiska system föreligger invecklade och differentierade relationer. En sådan mer skeptisk grundsyn bör ge upphov till nya och konstruktiva lösningar inom den tvåspråkiga lexikologin. (Jfr Snell-Hornby, 1990:209-226, Winter, 1992).

Noter

- 1 Detta förstudium ingår som ett delmoment i projektet "Kontrastiv valens - Ett lexikonorienterat studium i verbekvivalens" förlagt till Slaviska institutionen vid Göteborgs Universitet och utfört i samarbete med Inst. för svenska språket.
- 2 Med lemma menas en ordform som representerar ett ords samtliga böjningsformer.
- 3 I lemma-lexem-modellen kännetecknas artiklarnas huvudbetydelser, lexem, av "en väl identifierbar och relativt pregnant kärnbetydelse, kring vilken ett antal besläktade underordnade betydelser, kallade transfererade betydelser, kan gruppera sig." (Järborg, 1989:7-8.) Här signaleras lexemets underbetydelser på olika ställen i artikelns mikrostruktur, t ex i kommentardelen eller konstruktionsdelen.
- 4 Termerna 'referentiell' och 'relationell' har lånats från Norén (1990:20-22). Termen 'referentiella' använder Norén för ords egenbetydelse, usuella, betydelse, i kontrast till verbets relationella betydelse där man beaktar verbets egenskap att implicera vissa semantiska typer av fyllnadsled.
- 5 Termen fyllnadsled används här primärt för att hänvisa till det syntaktiska perspektivet i kontrast till 'argument' som framhäver det semantiska perspektivet. Sekundärt används den för att referera till valensled generellt.

Litteratur

- Allén, S. 1971. *Nusvensk frekvensordbok baserad på tidningstext. 1. Lemman*. Stockholm: Data linguistica 4.
- Allerton, D.J. 1982. *Valency and the English Verb*. London: Academic Press.
- Arnaud P.J.L. & Béjoint, H. (eds), 1992. *Vocabulary and Applied Linguistics*. Houndmills and London: Macmillan Academic and Professional Ltd.
- Brodda, B. 1973. (K)overta kasus i svenskan. (Pilus 18). Inst. för lingvistik. Stockholms universitet.
- Collins COBUILD *English Language Dictionary*. 1987. London and Glasgow.
- Hartmann, R.R.K. (ed.) 1984. *LEXeter '83 Proceedings*. Tübingen: Niemayer.
- Järborg, J. 1988. Towards a Formalized Lexicon of Swedish. I *Studies in ComputerAided Lexicology*. Stockholm: Almqvist & Wiksell International. s. 140-158.
- Järborg, J. 1989. Betydelseanalys och betydelsebeskrivning i Lexikalisk databas. Inst. för språkvetenskaplig databehandling. Göteborgs Universitet.
- Longman *Dictionary of Contemporary English*. (LDOCE). 1978. London.
- Norén, K. 1990. Svenska partikelverbs semantik. Institutionen för nordiska språk. Göteborgs universitet
- Snell-Hornby, M. 1992. Bilingual Lexicography. I Tomaszczyk et al. 1990:209-227.
- Somers, H.L. 1987. *Valency and case in computational linguistics*. (Edinburgh information technology series:3). Edinburgh University Press.
- Svensén, B. 1987. *Handbok i lexikografi*. Stockholm: Esselte Studium och Tekniska nomenklaturcentralen
- Svensk ordbok*. 1986. (Andra upplagan.) Utarbetad vid Språkdata. Göteborgs universitet. Stockholm. Esselte Studium.
- Tomaszczyk, J. & Lewandowska-Tomaszczyk, (eds) 1990. *Meaning and Lexicography*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Toporowska-Gronostaj, M. 1991. Mot ett formaliserat verbvalenslexikon. (Opublicerad licentiatuppsats.) Göteborgs universitet: Inst. för språkvetenskaplig databehandling.
- Winter, C. 1992. Bilingual Dictionaries: Between Language and Speech. In Arnaud P. J.L. & Béjoint, H. (eds): 1992:41-51.
- Zgusta, L. 1984. "Translational equivalence and the bilingual dictionary. In Hartman: 1984:147-154.
- Zgusta, L. 1987. "Translational Equivalence in a Bilingual Dictionary. Dictionaries. Number 9, 1987:1-47.

Betydningsbeskrivelser ved hjælp af genbrug og korpus

Ebba Hjorth og Knud Troels Thomsen

Den Danske Ordbog, der bliver på 6 bind, skal udkomme i tiden 1998-99. Den Danske Ordbog beskriver sproget efter 1950 med hovedvægt på perioden 1983-92¹.

To af ordene i dette foredrags titel markerer særlige forhold ved Den Danske Ordbog. Det er ordene **genbrug** og **tekstkorpus**. DDO opbygger et **tekstkorpus** på 40 millioner ord, som er det empiriske grundlag for ordbogens lingvistiske undersøgelser, og hvorfra eksempler til ordbogsartiklerne hentes. DDO's stramme tidsfrister gør, at **genbrug** af ordbogsdata bliver et nøglebegreb.

Ordbogsredaktører har altid stået på hinandens skuldre. Leksikografer har gennem tiderne i forskellig grad ladet sig inspirere af hinanden. Eller sagt på en anden måde: man har skrevet af efter hinanden i vekslende grad. Ofte har ordbogsredaktørers vanskeligste opgave været at formulere en betydningsbeskrivelse, der på én gang indeholdt **det samme** som inspiratorens betydningsbeskrivelse, sagde det på en **fikser** måde og **skjulte afhængigheden**.

I vores miljøbevidste tid er genbrug imidlertid blevet et respektabelt ord. Modviljen mod resursepild og kravet om hurtige resultater er også i leksikografien ligesom i den øvrige del af samfundslivet blevet hverdagssynspunkter, og dermed er genbrugstanken ved at blive sat i system.

I DDOs idégrundlag spiller genbrugstanken en afgørende rolle. Talrige steder i "Plan for en ordbog over dansk nutidssprog 18. december 1989"², som er vores arbejdsgrundlag, optræder ord og formuleringer, der taler om genbrug.

Nogle få eksempler: I omtalen af kilder til DDO hedder det (Plan etc. s.9):

“Til brug for redaktionsarbejdet fremskaffes i løbet af en forberedende fase et omfattende materiale bestående af nyexcerperede kilder og allerede tilgængelige citatsamlinger og ordbøger.”

Om disse ordbøger siges der lidt senere (Plan etc. s.14):

“Hovedparten af dem er tilgængelige i maskinlæsbar form, hvilket kan udnyttes til fremstilling af en slags "råmanuskript" ved delvis automatiske metoder.”

I planen gennemgås de enkelte oplysningstyper, som ønskes bragt ved hvert opslagsord, og det opregnes, hvilke leksikalske kilder der findes til de enkelte kilder. Under Semantik hedder det:

“- betydningsoptdeling/emneklassifikation: V&B, B&H, ...

- definition/parafrase: ODS, ODS-S ...”

Idéen var altså, at betydningsoptdelingen og emneklassifikationen og definitionerne skulle bygge på og lade sig inspirere af først og fremmest V&B, B&H og ODS. V&B og B&H er store oversættelsesordbøger, der er kommet i nye reviderede udgaver inden for det sidste par år. Begge ordbøger findes derfor ikke kun i maskinlæsbar form, men også i en struktureret og velbeskrevet form, der tillader, at der uden alt for store vanskeligheder uddrages bestemte oplysningstyper af de enkelte artikler, f.eks. betydningsoptdeling og emneklassifikation.

Tanken var derefter, at betydningsoptdelingerne i ODS skulle gøres maskinlæsbare, og at disse betydningsoptdelinger sammen med V&Bs og B&Hs betydningsoptdeling og emneklassifikation skulle sammenmanipuleres til et artikelskelet, et "råmanuskript" til brug for artikelskrivningen til DDO.

ODS er som de fleste vil vide et 28 bind stort værk, hvis første bind udkom i 1918 og hvis sidste ordbogsbind var på gaden i 1954. Ingen kan forvente en fast artikelstruktur i et sådant værk, og stillet over for de vanskeligheder, der er i at gøre så stort et værk maskinlæsbart og derefter at skulle strukturere oplysningerne, så det, der er relevant for DDO, udskilles, søgte vi nye veje.

Både V&B og B&H bygger i deres førsteudgaver på ODS. Og det var vores indtryk, at betydningsoptdelingerne i de to oversættelsesordbøger i lige så høj grad var en afspejling af ODS's betydningsoptdeling, som det var en betydningsoptdeling, der var rettet mod oversættelsesproblemer.

Lad os se på et par eksempler:

Betydningsoptdelingen for substantivet **nadver** ser i V&B således ud (de engelske ækivalenter er fjernet):

nadver 1 (*rel.*):
-en, den hellige ~, Herrens ~, -ens sakramente

2 (*glds. & rel.: aftensmad*)
den sidste ~; den store ~; deltage i -en

I B&H ser den således ud:

nadver 1 (*vx*)

2 (*relig*) (*den hellige ~*)

Den tilsvarende artikel i ODS ser ud som følger:

nadver **1** (.i rigsspr. nu bibl., arkais. ell. m.h.t. forhold paa landet) dagens sidste maaltid; aftensmaaltid; aftensmad.

2 (især bibl., kirk.) dels om det sidste paaskemaaltid, Jesus spiste sammen med sine disciple, skærtorsdag aften, dels om den derfra stammende kirkelige handling (sakramente), hvor præsten ved alteret uddeler (vin og) brød til menigheden efter fremsigelse af indstiftelsesordene; altergang; kommunion; undertiden ogs. om vinen og brødet, der uddeles.

Kombinerer vi de tre ordbøgers betydningsopdeling, ser vi, at der er god overensstemmelse mellem de tre ordbøger:

nadver **1** (rel.):
-en, den hellige ~, Herrens ~, -ens sakramente

2 (relig) (den hellige ~)

2 (især bibl., kirk.) dels om det sidste paaskemaaltid, Jesus spiste sammen med sine disciple, skærtorsdag aften, dels om den derfra stammende kirkelige handling (sakramente), hvor præsten ved alteret uddeler (vin og) brød til menigheden efter fremsigelse af indstiftelsesordene; altergang; kommunion; undertiden ogs. om vinen og brødet, der uddeles

2 (glds. & rel.: aftensmad)
den sidste ~; den store ~; deltage i -en

1 (vx)

1 (.i rigsspr. nu bibl., arkais. ell. m.h.t. forhold paa landet) dagens sidste maaltid; aftensmaaltid; aftensmad.

Jeg tillader mig hér at se bort fra, at man måske kan diskutere V&B's placering af ordforbindelserne under betydning 2.

Et andet eksempel:

Adjektivet **god** gives i V&B 6 engelske ækvivalenter.

god, adj.

- 1 good
- 2 (*rigelig*) good
- 3 (*stor, grundig etc.*) good
- 4 (*i god stand, ikke fordærvet, etc.*) good .. ; sound
- 5 (*gyldig, ægte*) good
- 6 (*artig*) good

De 6 ækvivalenter er forsynet med en kortfattet dansk betydningsbeskrivelse, og som det ses er oversættelsen i alle tilfælde *good*. Vi kan altså regne med, at betydningsopdelingen ikke er foretaget af hensyn til de engelske ækvivalenter, men er foretaget ud fra danske forhold.

Overensstemmelser som ved ordet *nadver* og struktureringer som ved ordet *god* har fået os til at afstå fra at gøre betydningsopdelingerne (og betydningsbeskrivelserne) fra ODS maskinlæsbare. Vi har i stedet for anvendt V&B og B&H til etableringen af en database, der indeholder betydningsopdelingen plus de faste forbindelser fra de to ordbøger. Det ser i øjeblikket således ud:

| | |
|--|---|
| <p>Artikel <i>nadiv</i></p> <p>Artikel</p> <p>Hoved</p> <p>Kilde R006</p> <p>OpsOrd <i>nadver</i></p> <p>Klasse sb</p> <p>Bøjn</p> <p>ental +en</p> <p>Semantik</p> <p>Betr</p> <p>Betbesk</p> <p>Formbrug</p> <p>Koll</p> <p>Tyd</p> <p>Tydspec</p> | <p><i>nadver</i> sb. +en.</p> <p>1. vx BH91</p> <p>2. relig BH91</p> <p>3. rel. VB90</p> <p>4. glds.& rel.: alt VB90</p> <p>a) <i>nadveren</i> VB90</p> <p>b) den hellige <i>nadver</i> BH91</p> <p>c) den hellige <i>nadver</i> VB90</p> <p>d) Herrens <i>nadver</i> VB90</p> <p>e) <i>nadverens</i> sakramente VB90</p> <p>f) den sidste <i>nadver</i> VB90</p> <p>g) den store <i>nadver</i> VB90</p> <p>h) <i>deltage i nadveren</i> VB90</p> |
|--|---|

Figur 1

Betydningsopdelingen fra de to ordbøger er anbragt efter hinanden, og derefter følger ordforbindelserne. Denne struktur vil nok blive ændret, så ordforbindelserne bliver anbragt under deres respektive betydninger, og muligvis vil også enslydende betydningsopdelinger fra de to ordbøger blive slået sammen.

Ebba Hjorth og Knud Troels Thomsen

Et sådant skelet, var det altså tanken, skulle konfronteres med vort 40 millioner ord store korpus. Dette var så at sige det ideologiske grundlag. Og hvad sagde så virkeligheden, da vi skulle til at fremstille artikler som f.eks. *nabo*.

Er ordbogsskeletter eller råartikler som disse:

| | |
|---------------------------------------|---|
| Kildev | |
| SkeletOpl | a) naboerne <i>BH91</i> |
| lykdoc | b) nabo i lejligheden ved siden af <i>BH91</i> |
| Citat naboerne | c) han har 10 km til nærmeste nabo <i>BH91</i> |
| Kilde | d) omgås med naboerne <i>BH91</i> |
| Kildek <i>BH91</i> | e) vor nabo mod syd (ogs) <i>BH91</i> |
| lykdoc | f) umiddelbar nabo <i>BH91</i> |
| Citat nabo i lejligheden ved siden af | g) være nabo til (= støde op til) <i>BH91</i> |
| Kilde | h) ingen har fred længere end hans nabo vil <i>VE90</i> |
| Kildek <i>BH91</i> | i) nærmeste nabo <i>VE90</i> |
| lykdoc | j) to miles til nærmeste nabo <i>VE90</i> |
| Citat han har 10 km til nærmeste nabo | k) optræde som en god nabo <i>VE90</i> |
| Kilde | l) have til nabo <i>VE90</i> |
| Kildek <i>BH91</i> | m) være nabo til <i>VE90</i> |
| lykdoc | |
| Citat omgås med naboerne | |
| Kilde | |
| Kildek <i>BH91</i> | |
| lykdoc | |
| Citat vor nabo mod syd (ogs) | |
| Kilde | |
| Kildek <i>BH91</i> | |
| lykdoc | |

Figur 2

et godt og sikkert grundlag at stå på, når man skal vende sig til sit 40 millioner ord store korpus og beskrive ordet *nabos* semantik, som i en meget foreløbig udgave kunne se således ud:

nabo *sb.*

...

1 *person, gruppe af personer eller virksomhed, som bor tættest på dig* inens ville hun gå over til naboen og rings til jordemoderen *Grethe Stenbæk Jensen: Ud af stedet (1987)*. Hun hørte lyden af toiletspapirruvlen blive drejet hos naboen *Vita Andersen: Hvat' for en hånd vil du ha' (1987)*.

2 *person, genstand eller bygning, som bor eller har fast opholdssted i nærheden af din bopæl* Trods sine næsten firs år havde hun vovet sig ned i byen efter en baguette i stedet for at appellere til en af sine nærmeste naboer *Erik Pouplier: Årets gang i Provence (1987)*.

3 *person eller genstand, som (tilfældigt) befinder sig i nærheden af dig* Min nærmeste nabo var en imponerende bronzemodel af Buckingham Palace placeret oven på en stor skibskiste *Erik Mortensen: 77001 nat (1990)*.

4 *land eller folk, som grænser op til et andet* vor nabo mod syd *Blinkenberg & Højbye (1991)*.

5 *et andet menneske, medmenneske* Den slags [*abort*] er kun noget man læser om. Det sker kun for naboen *Lilli Meldgaard Hestriksen: Ude og Hjemme (18/7 1990)*.

...

Figur 3

I det følgende behandles problemstillinger i forbindelse med genbrug af ordbøger ved nyredigering.

A. Oversættelsesordbøger da-fremmedsprog (OO) og B. Fremmede betydningsordbøger (FBO) som hjælpemiddel ved

1. lemmaselektion
 - a. generelt
 - b. ift. ODS
2. betydningsdifferentiering
 - a. generelt
 - b. ift. ODS

A: Oversættelsesordbøger da-fremmedsprog.

Ad 1a:

OO kan give inspiration til søgning i korpus mhp. at få udvalgt søgeord, især i form af faste forbindelser. Men der består dels en ikke ringe risiko for at få udvalgt ikke-existerende danske søgeord, da en del af lemmabestanden i OO ofte er mere eller mindre direkte omvendinger af målsprogsudtryk, som enten ikke er genuint danske og/eller er dårligt dansk, fx [omgås *med* naboerne] (B&H) (: at omgås naboerne), [umiddelbar nabo] (B&H) (: er ikke nogen fast forbindelse), [to miles til nærmeste nabo] (V&B) (: er ikke specielt dansk), dels en risiko for at møde en stor mængde støj i form af omvendinger, fx [forsyne er regiment med officerer = to officer a regiment] (V&B) (: er ikke noget begreb i DK), [statssekretær] (:findes ikke i DK) og forældede udtryk, fx [officerspatrulje] (V&B), [sødagtig] (B&H); både forældet og omvendning, fx [søindrullering, kontor for = quartier maritime] (Le grand Larousse, 1960, ['quartier']).

Disse typer søgeord vil for det første fra brugerens side derfor ikke blive betragtet som mulige søgeord, og er således i bedste fald redundante, for det andet ligge uden for DDOs sigte.

Ad 1b:

OO må i modsætning til ODS antages at være ført á jour til en vis grad, og vil derfor kunne fungere som kilder til valg af nye, ikke i ODS optagne ord og ordforbindelser, fx [hørebrille]; [skrælle en cykel]. - Som følge af ODS-baseringen vil der være en del støj i form af forældede ord. Men i modsætning til ODS vil enten OOs oversættelse af et givet ord og/eller OOs område-tilordning kunne fungere som indikator for, om et gammelt eller ældre ord har fået en ny betydning, fx [skærm (TV)] eller om der er indlånt nye homografer, fx [fil (edb)].

Tages i betragtning, at korpus, skulle det være nok så repræsentativt, altid vil være underlagt en vis grad af tilfældighed, kan der være væsentlige ord- og kollokationsforekomster, der ikke er repræsenteret der. Set i det perspektiv kan der således være god me-

ning i at inddrage de elektronisk læsbare OO, idet de netop her kan få funktion som 'the missing link' mellem vort korpus og ODS, og således bidrage til at bøde på tilfældighedsfaktoren, ligesom de formentlig kan supplere Sprognævnets liste over nye ord.

Ad 2a:

OO er i et vist omfang anvendelige til betydnings*differentiering*, men kun i meget begrænset omfang til betydnings*beskrivelse*. Deres sigte er mere rettet mod oversættelsesproblemer, som ofte bunder i uskarpe betydnings*differentieringer*, end mod egentlig betydnings*beskrivelse*. Ikke underligt, for OO forudsætter *modersmålskompetence*, og derfor er de ikke tvunget til at beskrive modersmålet, medmindre det er relevant for oversættelsesformålet.

Således har de to oversættelsesordbøger eksempelvis kun optaget én betydning af nabo, hvor vi, DDO, har defineret fem. Dette formentlig fordi ordet nabo i alle betydninger oversættes til det samme. Til gengæld kan betydnings*differentieringerne* i oversættelsen give anledning til at overveje, hvorvidt der bør foretages yderligere *differentieringer* eller i det mindste at undersøge, om det er fornuftigt, dvs., om der er belæg for at foretage en lignende *differentiering*, fx [sheet = kortblad] (som ikke findes i ODS).

Ad 2b:

ODS' betydningsstruktur er overtaget ukritisk, men i forhold til ODS medtages i reglen betydeligt færre betydninger, idet flere betydninger ikke nødvendigvis resulterer i forskellige oversættelser, hvorfor det ville være redundant at *differentiere* af oversættelsesenssyn. Til gengæld kan der, som oven anført, være angivet nye *differentieringer*, lige som gamle kan være gledet ud af sproget. Se i øvrigt ovenfor.

B. Fremmede betydningsordbøger.

Ad 1:

Bortfalder iflg. sagens natur

Ad 2a og b:

For at nyttiggøre FBO forudsættes naturligvis en relevant fremmedsprogskompetence. Generelt må man gå ud fra, at der i forhold til danske beskrivelser, måske bortset fra visse internationalismer, kan og vil forekomme større eller mindre geografi- og kulturafhængige forskelle i betydnings*beskrivelsen*. Der ville således være meget stof til konflikt, hvis den Den Danske Ordbog skrev af efter FBO. Når vi på DDO beskriver ordbetydninger, har vi derfor indført den procedure, at vi først danner os et indtryk af betydningen ved introspektion (på grundlag af exx fra korpus), derefter konsulterer vi ODS, hvor det er muligt, og først i tredje instans skæver vi til FBO og OO.

For et få et så stabilt indtryk som muligt af et ords betydning(er), udarbejdes en oversigt over betydningsoplysninger på grundlag af samtlige eksempler fra de værker, der øn-

skes inddraget, idet betydningsoplysningerne typologiseres. Typologiseringen er ikke i sig selv noget sandhedsvidne, da den ifølge sagens natur er en arbitrær størrelse. Den tjener da også alene det formål, at søge at skabe et sammenligningsgrundlag de enkelte værker imellem. Nedenfor bringes exx på, hvorledes disse oversigter ser ud i praksis, og de er konstrueret til de to lemmata [*bjerg*] og [*nabo*]. Da oversigterne taler for sig selv, vil de ikke blive kommenteret yderligere.

bjerg:

ODS:

[**Bjerg**][1] større forhøjning i jordens .. overflade og 1.2) om forhold uden for Danmark (og paa Bornholm): stor (kuplet eller tilspidset) forhøjning især bestaaende af stenmasser]

| | |
|---------|----------------------|
| hvad | +forhøjning |
| hvor 1 | +i jordens overflade |
| hvor 2 | - |
| form 1 | +kuplet eller |
| form 2 | +tilspidset |
| form 3 | - |
| mater. | +stenmasser |
| størr.1 | +stor |
| størr.2 | - |
| smlgn.1 | - |
| smlgn.2 | - |
| gener. | - |
| antonym | - |
| encyc. | - |

AHD:

[**mountain**][1] A natural elevation of the earth's surface having considerable mass, generally steep sides, and a height greater than that of a hill]

| | |
|---------|-------------------------------------|
| hvad | +elevation |
| hvor 1 | +earth's surface |
| hvor 2 | - |
| form 1 | - |
| form 2 | - |
| form 3 | +steep sides, |
| mater. | - |
| størr.1 | - |
| størr.2 | +considerable mass |
| smlgn.1 | +height greater than that of a hill |
| smlgn.2 | - |
| gener. | +natural |
| antonym | - |
| encyc. | - |

DUDEN-U:

[**Berg**][Größere Erhebung im Gelände]

| | |
|---------|-------------|
| hvad | +Erhebung |
| hvor 1 | +im Gelände |
| hvor 2 | - |
| form 1 | - |
| form 2 | - |
| form 3 | - |
| mat. | - |
| stør. 1 | +größere |
| stør. 2 | - |
| smlgn.1 | - |
| smlgn.2 | - |
| gener. | - |
| antonym | - |
| encyc. | - |

OALD:

[**mountain**] [1] mass of very high rock going up to a peak]

| | |
|----------|-------------|
| hvad | +mass |
| hvor 1 | - |
| hvor 2 | - |
| form 1 | - |
| form 2 | + peak |
| form 3 | + very high |
| størr. 1 | - |
| størr. 2 | - |
| mater. | + rock |
| smlgn. 1 | - |
| smlgn. 2 | - |
| gener. | - |
| antonym | - |
| encyc. | - |

SAOB:

[BERG] [1] (ansenlig) opphøjning på jordens (1. annan fast himlakrops) yta, ofta med tanka på att upphøjningen hufvudsakligen består af stenmassa]

| | |
|----------|--------------------------------|
| hvad | + opphøjning |
| hvor 1 | + jordens yta |
| hvor 2 | + 1. annan fast himlakrops yta |
| form 1 | - |
| form 2 | - |
| form 3 | - |
| størr. 1 | + ansenlig |
| størr. 2 | - |
| mater. | + stenmassa |
| smlgn. 1 | - |
| smlgn. 2 | - |
| gener. | - |
| antonym | - |
| encyc. | - |

WDG:

[Berg] [1] Große, aus dem Gelände herausragende Erhebung, Ggs. Tal]

| | |
|----------|------------|
| hvad | + Erhebung |
| hvor 1 | + Gelände |
| hvor 2 | - |
| form 1 | - |
| form 2 | - |
| form 3 | - |
| mater. | - |
| stør. 1 | + groß |
| stør. 2 | - |
| mater. | - |
| smlgn. 1 | - |
| smlgn. 2 | - |
| gener. | - |
| antonym | + Tal |
| encyc. | - |

COBUILD:

[mountain] [1] A mountain is a very large raised part of the earth's surface with steep sides which are usually difficult to climb. A mountain is bigger and higher than a hill]

| | |
|----------|--|
| hvad | + raised part (of the earth's surface) |
| hvor 1 | + earth's surface |
| hvor 2 | - |
| form 1 | - |
| form 2 | + higher |
| form 3 | + steep sides |
| størr. 1 | + very large |
| størr. 2 | - |
| mater. | - |
| smlgn.1 | + higher |
| smlgn.2 | + bigger |
| gener. | - |
| antonym | - |
| encyc. | + usually difficult to climb |

nabo:

ODS:

[nabo] [1](hver af) de personer, der (i forsk. retninger) bor nærmest ved siden af en given person, ell. hvis ejendomme, jorder støder op til denne persons; mht. bebyggelse langs veje, gader: person, der bebor huset, ejendommen ved siden af (mods. Genbo); ogs. om (hver af) to i samme hus ell. samme lejlighed boende personer, hvis lejligheder eller værelser støder op til hinanden, betragtet i forhold til hinanden]

person 1 + person
 person 2 + given person
 persongrp. -
 lokalitet 1 -
 lokalitet 2 -
 genstand 1 -
 genstand 2 -
 fast bopæl + bo, bebo
 tilf. ophold -
 relationer
 hvem: person 1 og person 2
 hvad: bopæl eller ejendom
 hvorledes: støder op til hinanden
 aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret -
 antonym + genbo

ODS:

[nabo] [2] Videre anvendelse af bet. 1. 2.1) (især skr.spr.) om (folk, nation i et) tilstødende omraade ell. land.]

person 1 -
 person 2 -
 persongrp. 1 + folk, nation
 persongrp. 2 + folk, nation
 lokalitet 1 + tilstødende land el. område
 lokalitet 2 + tilstødende land el. område
 genstand 1 -
 genstand 2 -
 fast bopæl (+)
 tilf. ophold -
 relationer -
 hvem: persongrp. 1 og persongrp. 2
 hvad: område el. land
 hvorledes: støder op til hinanden
 aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret + især skr.spr.
 antonym -

ODS:

[nabo] [2.2] Om tilgrænsende ejendom ell. virksomhed]

person 1 -
 person 2 -
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 -
 lokalitet 2 -
 genstand 1 + ejendom ell. virksomhed
 genstand 2 + ikke defineret
 fast bopæl (+)
 tilf. ophold -
 relationer -
 hvem: -
 hvad: ejendom ell. virksomhed og en anden genstand
 hvorledes: støder op til hinanden
 aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret -
 antonym -

ODS:

[nabo] [2.3] (især skr.spr.) om person, der har (sidde-, ligge) plads ved siden af en anden]

person 1 + person
 person 2 + en anden
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 -
 lokalitet 2 -
 genstand 1 -
 genstand 2 -
 fast bopæl -
 tilf. ophold + sidde- eller liggeplads
 relationer
 hvem: person 1 og person 2
 hvad: plads (placering)
 hvorledes: ved siden af hinanden
 aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret + skr.spr.
 antonym -

DUDEN-U:

[Nachbar] [a] jmd., der in jmds. [unmittelbarer] Nähe wohnt, dessen Haus, Grundstück in der [unmittelbaren] Nähe von jmds. Haus, Grundstück liegt.]

person 1 + jmd.
 person 2 + jmds.
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 -
 lokalitet 2 -
 genstand 1 -
 genstand 2 -
 fast bopæl + wohnen
 tilf. ophold -
 relationer

hvem: person 1 og person 2
 hvad: Haus, Grundstück
 hvorledes: in der unmittelbaren Nähe

aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret -
 antonym -

DUDEN-U:

[Nachbar] [b] jmd., der sich in jmds. [unmittelbarer] Nähe befindet ... (Nachbarländer) ... (Nachbar)]

person 1 + jmd.
 person 2 + jmds.
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 + Nachbarland
 lokalitet 2 + Nachbarland
 genstand 1 -
 genstand 2 -
 fast bopæl -
 tilf. ophold + sich befinden
 relationer

hvem: person 1 og person 2 eller
 hvad: Nachbarland 1 og Nachbarland 2
 hvorledes: in der unmittelbaren Nähe

aktivitet -
 tiltaleform + Nachbar
 fagsprog -
 stilmarkeret -
 antonym -

AHD:

[neighbor] [1] One who lives near or next to another]

person 1 + one
 person 2 + another
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 -
 lokalitet 2 -
 genstand 1 -
 genstand 2 -
 fast bopæl + lives
 tilf. ophold -
 relationer

hvem: person 1 og person 2
 hvad: -
 hvorledes: near or next to

aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret -
 antonym -

AHD:

[neighbor] [2] A person, place or thing adjacent to or located near another]

person 1 + A person
 person 2 + another
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 + place
 lokalitet 2 + another (place)
 genstand 1 + thing
 genstand 2 + another thing
 fast bopæl ± adjacent to or located
 tilf. ophold ± adjacent to or located
 relationer

hvem: person 1 og person 2 (eller)
 hvad: lokalitet 1 og lokalitet 2 (eller)
 hvad: genstand 1 og genstand 2
 hvorledes: near another

aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret -
 antonym -

AHD:

[neighbor] [3] A fellow human being]

person 1 + human being
 person 2 ± (fellow)
 persongrp. 1 ± fellow
 persongrp. 2 -
 lokalitet 1 -
 lokalitet 2 -
 genstand 1 -
 genstand 2 -
 fast bopæl -
 tilf. ophold -
 relationer
 hvem: person 1 og person 2 eller persongrp. 1
 hvad: -
 hvorledes: fellow(-ship)
 aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret -
 antonym -

AHD:

[neighbor] [4] Used as a form of familiar address]

person 1 + (person)
 person 2 + (person)
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 -
 lokalitet 2 -
 genstand 1 -
 genstand 2 -
 fast bopæl -
 tilf. ophold -
 relationer
 hvem: person 1 og person 2
 hvad: -
 hvorledes: a form of familiar address
 aktivitet -
 tiltaleform + address
 fagsprog -
 stilmarkeret -
 antonym -

OALD:

[neighbour] [1a] person living next to or near another]

person 1 + person
 person 2 + another
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 -
 lokalitet 2 -
 genstand 1 -
 genstand 2 -
 fast bopæl + living
 tilf. ophold -
 relationer
 hvem: person 1 og person 2
 hvad: -
 hvorledes: next to or near
 aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret -
 antonym -

OALD:

[neighbour] [1b] person, thing or country that is next to or near another]

person 1 + person
 person 2 + another
 persongrp. 1 -
 persongrp. 2 -
 lokalitet 1 + country
 lokalitet 2 + another (country)
 genstand 1 + thing
 genstand 2 + another (thing)
 fast bopæl ± is
 tilf. ophold ± is
 relationer
 hvem: person 1 og person 2 (eller)
 hvad: lokalitet 1 og lokalitet 2 (eller)
 hvad: genstand 1 og genstand 2
 hvorledes: next to or near another
 aktivitet -
 tiltaleform -
 fagsprog -
 stilmarkeret -
 antonym -

Ebba Hjorth og Knud Troels Thomsen

OALD:

[neighbour] [2] fellow human being]

| | |
|---|---------------|
| person 1 | + human being |
| person 2 | ± (fellow) |
| persongrp. 1 | ± fellow |
| persongrp. 2 | - |
| lokalitet 1 | - |
| lokalitet 2 | - |
| genstand 1 | - |
| genstand 2 | - |
| fast bopæl | - |
| tilf. ophold | - |
| relationer | |
| hvem: person 1 og person 2 eller persongrp. 1 | |
| hvad: - | |
| hvorledes: fellow = samme slags | |
| aktivitet | - |
| tiltaleform | - |
| fagsprog | - |
| stilmærket | - |
| antonym | - |

SAOB:

[NABO] [b] jur. om ägare av fastighet i stad i förh. till ägare av fastighet som gränsar omedelbart intill den förra]

| | |
|--|----------------------------------|
| person 1 | + ägare |
| person 2 | ± ägare |
| persongrp. 1 | - |
| persongrp. 2 | - |
| lokalitet 1 | - |
| lokalitet 2 | - |
| genstand 1 | + fastighet |
| genstand 2 | + fastighet som gränsar ... |
| fast bopæl | ± (mulighed, men ikke defineret) |
| tilf. ophold | - |
| relationer | |
| hvem: person 1 og person 2 | |
| hvad: fastighet 1 og fastighet 2 | |
| hvorledes: som gränsar omedelbart intill den förra | |
| aktivitet | - |
| tiltaleform | - |
| fagsprog | + juridisk |
| stilmærket | - |
| antonym | - |

SAOB:

[NABO] [c] i utvidgad anv., om samhälle o. dyl. l. land l. folk (l. dess överhuvud) i förh. till angränsande samhälle]

| | |
|---|------------------------------|
| person 1 | (+ överhuvud) |
| person 2 | - |
| persongrp. 1 | + folk |
| persongrp. 2 | + angränsande folk |
| lokalitet 1 | + land, samhälle |
| lokalitet 2 | + angränsande land, samhälle |
| genstand 1 | - |
| genstand 2 | - |
| fast bopæl | + angränsande land |
| tilf. ophold | + (mulighed, ikke defineret) |
| relationer | |
| hvem: persongrp. 1 og persongruppe 2 (eller) hvad: lokalitet 1 og lokalitet 2 | |
| hvorledes: som er 'angränsande' | |
| aktivitet | - |
| tiltaleform | - |
| fagsprog | - |
| stilmærket | - |
| antonym | - |

SAOB:

[NABO] [d]† om person l. folk i förh. till ett naturföremål o.d. som är beläget i närheten av hans (dess) hemvist]

| | |
|---|----------------|
| person 1 | + person |
| person 2 | - |
| persongrp. 1 | + folk |
| persongrp. 2 | - |
| lokalitet 1 | + hemvist |
| lokalitet 2 | - |
| genstand 1 | + naturföremål |
| genstand 2 | - |
| fast bopæl | + hemvist |
| tilf. ophold | - |
| relationer | |
| hvem: person el. persongrp (og) hvad: lokalitet 2 | |
| hvorledes: som är beläget i närheten af lokalitet 1 | |
| aktivitet | - |
| tiltaleform | - |
| fagsprog | - |
| stilmærket | + forældet |
| antonym | - |

WDG:

[Nachbar] [jmd., der (unmittelbar) neben jmdm. wohnt, innerhalb einer Reihe neben jmdm. sitzt, steht ... (die im Osten an unseren Staat angrenzenden Staaten) .../umg.scherzh./ Anrede an den Nachbarn; /umg.veraltend/ wohlwollende Anrede an einen an einen Unbekannten;/ spött.]

| | |
|--------------|---|
| person 1 | + jmd. |
| person 2 | + jmdm./ ein(e) Unbekannte(r) |
| persongrp. 1 | - |
| persongrp. 2 | - |
| lokalitet 1 | + unser Staat |
| lokalitet 2 | + angrenzender Staat |
| genstand 1 | - |
| genstand 2 | - |
| fast bopæl | ± wohnen |
| tilf. ophold | ± sitzen, stehen |
| relationer | |
| hvem: | person 1 og person 2 (eller) |
| hvad: | unser Staat og angrenzender Staat (eller) : Anrede |
| hvorledes: | unmittelbar neben, angrenzend; |
| hvis | Anrede: wohlwollend eller spöttisch |
| aktivitet | - |
| tiltaleform | ± Anrede (hvis personer) |
| fagsprog | - |
| stilmarkeret | ± spött. |
| antonym | - |

COBUILD:

[neighbour] [1] A neighbour is 1.1) someone who lives near you]

| | |
|--------------|----------------------|
| person 1 | + someone |
| person 2 | + you |
| persongrp. 1 | - |
| persongrp. 2 | - |
| lokalitet 1 | - |
| lokalitet 2 | - |
| genstand 1 | - |
| genstand 2 | - |
| fast bopæl | + lives |
| tilf. ophold | - |
| relationer | |
| hvem: | person 1 og person 2 |
| hvad: | |
| hvorledes: | near (hinanden) |
| aktivitet | - |
| tiltaleform | - |
| fagsprog | - |
| stilmarkeret | - |
| antonym | - |

COBUILD:

[neighbour] [1] A neighbour is 1.2) someone who is standing or sitting next to you or something which is next to something else of the samekind]

| | |
|--------------|--|
| person 1 | + someone |
| person 2 | + you |
| persongrp. 1 | ± something |
| persongrp. 2 | ± something else |
| lokalitet 1 | ± something |
| lokalitet 2 | ± something else |
| genstand 1 | ± something |
| genstand 2 | ± something else |
| fast bopæl | - |
| tilf. ophold | ± uvist |
| relationer | |
| hvem: | person 1 og person 2 |
| hvad: | persongrp. 1 og persongruppe 2 (eller) lokalitet 1 og lokalitet 2 (eller) genstand 1 og genstand 2 |
| hvorledes: | sitting eller standing eller being next to (hinanden) |
| aktivitet | - |
| tiltaleform | - |
| fagsprog | - |
| stilmarkeret | - |
| antonym | - |

COBUILD:

[neighbour] [2] The neighbour of a country, county, state, etc is another country, county, state that is next to it]

| | |
|--------------|----------------------------|
| person 1 | - |
| person 2 | - |
| persongrp. 1 | - |
| persongrp. 2 | - |
| lokalitet 1 | + country, county, state |
| lokalitet 2 | + country, county, state |
| genstand 1 | - |
| genstand 2 | - |
| fast bopæl | + (ingen anden mulighed) |
| tilf. ophold | - |
| relationer | |
| hvem: | - |
| hvad: | lokalitet 1 og lokalitet 2 |
| hvorledes: | being next to each other |
| aktivitet | - |
| tiltaleform | - |
| fagsprog | - |
| stilmarkeret | - |
| antonym | - |

COBUILD:

[neighbour] [3] Your neighbour is a person who you have dealings with; an old-fashioned use]

person 1 + a person
person 2 + you
persongrp. 1 -
persongrp. 2 -
lokalitet 1 -
lokalitet 2 -
genstand 1 -
genstand 2 -
fast bopæl -
tilf. ophold ± ikke defineret
relationeer
 hvem: person 1 og person 2
 hvad: -
 hvorledes: dealing with (hinanden)
aktivitet -
tiltaleform -
fagsprog -
stilmarkeret + old-fashioned
antonym -

Noter

- 1 For nærmere oplysninger om DDO se Kjeld Kristensen: Den Danske Ordbogs tekst-korpus og SpORDhunde.
- 2 Dansk Ordbog. Plan for en ordbog over dansk nutidssprog. 18. december 1989. Af Ebba Hjorth, Iver Kjær, Ole Norling-Christensen og Hanne Ruus.

Forkortelser

| | |
|---------|--|
| AHD | The American Heritage Dictionary of the English Language. Third Edition. Boston 1992. |
| B&H | Blinkenberg, Andreas & Poul Høybye: Dictionnaire Danois-Français / Dansk-fransk Ordbog. 4. udg. ved Jens Rasmussen & al. Bind 1-2. København 1991. |
| COBUILD | Collins Cobuild English Language Dictionary. London 1987. |
| DDO | Den Danske Ordbog. Under udgivelse af Det danske Sprog- og Litteraturselskab. København. |
| DUDEN-U | Duden deutsches Universalwörterbuch. Mannheim 1989. |
| OALD | Oxford Advanced Learner's Dictionary of Current English, Encyclopedic Edition. Oxford 1992. |
| ODS | Ordbog over det Danske Sprog. Udgivet af Det danske Sprog- og Litteraturselskab. Bind 1-28. København 1918-56. |
| SAOB | Ordbok över Svenska Språket utgiven av Svenska Akademien. Band 1ff. Lund 1898ff. |
| V&B | Vinterberg, Hermann & C.A. Bodelsen: Dansk-engelsk Ordbog. 3. udg. ved Viggo Hjørnager Pedersen. København 1990. |
| WDG | Wörterbuch der deutschen Gegenwartssprache. Herausgegeben von Ruth Klappenbach und Wolfgang Steinitz. Band 1-6. Berlin 1964-1977. |

Nogle generelle problemer i forbindelse med udarbejdelse af bilinguale ordbøger for ældre stadier

Karl Hyldgaard-Jensen

Når man udarbejder ordbøger for to nutidssprog, er der jo ingen der er i tvivl om at vi har med to selvstændige og jævnbyrdige sprog at gøre. Men når man sammenligner ældre sprogstadier, må man først gøre sig klart, hvilken status disse stadier har. Er de egentlige diasystemer eller udgøres de bare af en samling skriftdialekter?

For at et kommunikationsinstrument skal kunne betegnes som et sprog kræves der to ting:

1) Sprogets system skal være en fællesnævner for de mange subsystemer der findes inden for det område hvor sproget bruges. Sproget skal være en overbygningsform, der f. eks. gælder som fælles kultursprog eller skriftsprog.

2) For det andet skal sprogets system adskille sig på en markant måde fra de omgivende nabosprogs systemer.

ad 1) Sprogets system: Takket være Hansaen opstod der i middelalderen et middelnedertysk skriftsprog, der brugtes i korrespondancer mellem Hansestæderne i Nordtyskland, Baltikum, Skandinavien, Holland, Belgien, England. Dette skriftsprog havde skriftsproget i den Hansestad der efterhånden overtog føringen efter Visby, nemlig Lübeck, som basis, d.v.s. det var en nordnedertysk dialekt. Der fandtes altså efter ca. 1350 et fælles forvaltningsskriftsprog, omend med lokale afvigelser. Det er uklart om der også fandtes et fælles mundtligt sprog. Men dette middelnedertyske skriftsprog forsvandt igen, da Hansaens magtposition reduceredes omkring 1500 på grund af Amerikas opdagelse, hvorved handelen blev koncentreret til Sydvesteuropa i stedet for til Østersøområdet. Dødsstødet gav dog Reformationen, der gennemtrufede en blanding af mellemtyisk og sydtyisk som skriftsprog svarende til Luthers østmellemtyske sprog. Nynedertysk (altså plattysk) er ikke noget sprog, kun en samling dialekter uden fælles norm.

ad 2) Afgrænsningen af *middelnedertysk* mod syd, der sprogligt sker ved hjælp af den højtyske konsonantforskydning, falder også sammen men Hansaens udbredelse mod syd i mellemtyiskland. Dog sker der i middelalderens forløb en forskydning i brugen af middelnedertysk skriftsprog mod nord. F. eks. fortrængtes middelnedertysk allerede i

1417 af middelhøjtysk i Halles cancelli (som folkesprog dog betydeligt senere). Hele området Halle, Wittenberg, Merseburg blir tidligt høj- tysk. Omkring 1550 ophører det fælles m.nedertyske skriftsprog. Vi kan altså kun for perioden 1350 til 1550 hævde at middelnedertysk var et sprog.

Når det gælder *gammeldansk*, så har det jo altid været betragtet som et sprog. Der findes et forstadium til gammeldansk og en fortsættelse op til nutiden. Gammeldansk udgør altså et diasystem. Der fandtes også klare tendenser til en fælles norm udgående fra de kongelige cancellier og der fandtes et dansk kultursprog ved sidan af latin.

Afgrænsningen mod middelnedertysk i syd er sprogligt også klar, selv om der nu gøres forsøg på at vise at de to sprog typologisk set er meget ensartede; Lena Moberg hævder (Lågtyskt och svenskt i Stockholms medeltida tänkeböcker, 1989) at den store lighed mellem systemerne i middelnedertysk og fornsvenskan gjorde at et blandingssprog kunne opstå i de svenske hansestæder. Afgrænsningen mod de øvrige nordiske sprog regner man jo som et faktum allerede i tidlig middelalder.

Der er altså tale om en sammenligning mellem to sprog inden for en given periode: Senmiddelalderen.

Benyttede corpora:

I. Mnt.: For middelnedertysks vedkommende benyttes Mnd. Wörterbucharchiv, Hamburg med publikationen: Lasch-Borchling-Cordes "Mittelniederdeutsches Handwörterbuch". Ordbogen er ikke fuldstændig; sidste bind er s-, men der er huller i alfabetet. Den fuldstændige ordbog ved Schiller-Lübben "Mittelniederdeutsches Wörterbuch" kan kun bruges med stor forsigtighed. Schiller og Lübben var begge jurister uden filologisk uddannelse; de medtar ukritisk mellemtyiske ord og ordformer som nedertyske fra de områder, hvor nedertysk grænser til østmellemtysk. For det mnt. lovsprogs vedkommende benytter jeg mine egne EDB-tekster til samtlige nordtyske stadsretter, Sachsen-spiegel og andre juridiske tekster skrevet på middelnedertysk. Som sammenligningsmateriale fremstillede jeg EDB-tekster til nogle mht. og mellemtyiske stadsretter. Konkordanser, ordlister og frekvenslister hertil optager det middelnedertyske ordforråd inden for lovsproget 100%. Arbejdet udførtes 1969-70, altså for snart 25 år siden. D

II. Det ville have været til stor hjælp, hvis vi også havde haft et tilsvarende EDB-corpus for de danske landskabs- og købstadslove fra middelalderen. Jeg har selv måttet excerperer disse tekster; men nu findes der jo arkiv- samlinger + database til Gammeldansk ordbog på Københavns universitet, hvor jeg formoder at lovene ud over diplomerne er excerperet in extenso. Dog benytter jeg også Otto Kalkars Ordbog til det ældre danske sprog (1881-1918) og for lovsprogets vedkommende specielt G.F.V. Lund: Det ældste danske skriftsprogs ordforråd. Ordbog til de gamle danske landskabslove, de sønderjyske stadsretter samt øvrige sprogmindesmærker (fra omtr. 1200-1300), (1877). Endvidere har man god hjælp af Peter Skautrupps ordliste til hans udgave af Jyske Lov 1941.

Da det er af sproghistorisk interesse at konstatere, hvordan de middelnedertyske lå-

neord i de skandinaviske landes middelalderlige forvaltningssprog er trængt ind, sammenligner jeg desuden gennemgående med de aktuelle ords forekomst i fornsvenskan. Som corpora benyttes her K. F. Söderwall: Ordbok öfver svenska medeltids-språket 1884-1918 + supplement 1925-1973 samt C. J. Schlyter: Corpus iuris sueco-gothorum antiqui, 1827-77. Endvidere har jeg konsulteret databaser for fornsvenskan som findes ved universiteterne i Göteborg og Lund (SAOB).

Når man skal sammenligne to sprogs ordforråd, må man have et sammenligningsgrundlag, det såkaldte *tertium comparationis*. Jeg arbejdede først med et system, hvor jeg valgte den latinske terminus som sammenligningsgrundlag. Men det er langt fra alle middelnedertyske og gammeldanske retstermer, der er belagt i latinske versioner af de pågældende love; og det er heller ikke givet at det latinske ord dækker over det samme begreb i de to sprogs lovtekster.

Nogle tyske ordbøger, der sammenligner de to former af tysk: højtysk og nedertysk, har ikke noget *tertium comparationis*; de vælger i stedet at udgå fra det ene af sprogene, nemlig højtysk som basis. Resultatet er mærkeligt. Når der dukker et nedertysk ord op i den alfabetisk ordnede fremstilling, som ikke har nogen pendant i højtysk, laver man det nedertyske ord om, så det får højtysk form, f. eks. skrives mnt. *bespråken* 'anklage' som lemma med en asterisk + *beisprachen*, en form der ellers ikke forekommer i højtysk. For tyskere er det naturligt at vælge den konstruerede højtyske form som lemma, da de fleste ikke kan nedertysk, og desuden er metasproget i ordbogsartiklerne jo højtysk. Denne fremgangsmåde anvender bl.a. "Deutsches Rechtswörterbuch", Heidelberg og "Wörterbuch der mhd. Urkundensprache bis zum Jahr 1300", Berlin. F. eks. står der som lemma her en konstrueret mht. form **bierphenninc* for det mnt. ord *bêrpenmink* 'lidkøb'. Sommetider står det mnt. ord efter det konstruerede højtyske, adskilt ved skråstreg. Schiller-Lübbers og Lasch-Borchling-Cordes' mnt. ordbøger udgår derimod konsekvent fra den nedertyske form som lemma, selv om metasproget er højtysk; f. eks. har de som lemma *anevangen* i stedet for *anfangen* (mht. *anvâhen*).

I de artikler jeg udarbejder nu, har jeg i stedet valgt at udgå fra det juridiske begrebssystem, som de middelnedertyske og gammeldanske termer indgår i. Der tages her højde for om det pågældende retsinstitut er et oprindeligt germansk begreb eller om der foreligger påvirkning fra kanonisk og romersk ret.

Jeg spørger altså: Hvordan betegnes det aktuelle begreb (retsinstitut) i middelnedertysk og i gammeldansk? Man kan da umiddelbart afgøre, om betegnelserne har den samme plads i begrebssystemet eller om de ikke dækker hinanden. Desuden kan eventuelle synonyme let indplaceres i systemet. Hvilket sprog man vælger til fremstilling af begrebssystemet kan være ligegyldigt. Hvis metasproget er højtysk, er det naturligt at vælge højtysk.

Når der er flere termer (altså ikke varianter) for samme retsinstitut, er det vanskeligt at afgøre, hvilken man skal vælge som lemma (ty. Stichwort). I det tilfælde jeg illustrerer i fig. 1, har "Handwörterbuch zur deutschen Rechtsgeschichte" valgt den højtyske form af den juridiske term der var mest almindelig i det 15.-16. århundrede, fordi der er tale

om en fremstilling af tyske forhold i middelalderen, d.v.s. "Kummer". Jeg har i stedet valgt "Beschlagnahme og Arrest" der som metasprogets termer dækker både nordiske og tyske forhold. Ved valget af "Kummer" som lemma synes man at være udgået fra "Frühneuhochdeutsches Wörterbuch" ved Anderson-Goebel-Reichmann (1986ff.). Oskar Reichmann skriver her i Bd. 1, S.17, at lemmaet skal ansættes på basis af alle varianter leksikografen på grundlag af sit intuitiv-linguale vidende konstituerer som sådanne, d.v.s. intuitivt opfatter som varianter af samme ord. Dog skal den ideale form der vælges ikke være en lydlovmæssig afledt form; i stedet skal man vælge den form der bruges eller ville være blevet brugt af de sociologisk ledende skrivergrupper og trykkerier i litterære og oplysende tekstsorter fra den givne periode, altså en idealiseret form. "Frühneuhochdeutsches Wörterbuch" ansætter derfor i position 1 kun eet ord (een ordform) som lemma. I position 2 opføres et udvalg af varianter med angivelse af lokalisering, datering og sociolekt. Dette sker også intuitivt. I position 3 giver man morfologiske data, i position 4: betydningsangivelser. I position 5 angives den geografisk-kronologiske udbredelse, i position 6 belæg og kildehenvisninger. Position 7 optages af det onomasiologiske felt (=synonymer), mens de syntagmer ordet indgår i anbringes under position 8. Eftersom Reichmanns teoretisk baserede opbygning af ordartikler for ældre sprogstadier må anses som "le dernier cri" inden for denne type af leksikografi, tar jeg hensyn til hans forskellige punkter ved opbygningen af mine artikler, men vælger som sagt en helt anden fremgangsmåde end hans alfabetisk baserede.

Både gammeldansk/fornsvenskan og middelnedertysk savner en fast fælles norm i ortografien. Ved normalisering og ansættelse af lemma må vi derfor agere på samme måde som Reichmann, nemlig vælge den form som vi intuitivt på basis af belæggene anser er den mest centrale i en given periode. Som man vil se af fig. 1, har jeg ordnet de mnt. termer *kronologisk* med sidestillede varianter. For gammeldansk og fornsvenskans vedkommende skelner jeg dog mellem arveord og låneord. Jeg mener det er berettiget at dele synonymerne op på denne måde, selv om brugen af den ene kategori måske ikke er så udbredt i den givne periode. Semikolon angiver grænsen mellem arveord og låneord i glda. og fsv.

Arveordene er i fig. 1-eksemplet fælles for middelnedertysk og de skandinaviske sprog (glda. *vædhæ*, *opholdæ*, *hindræ* og de fsv. paralleller; *hindræ* er eventuelt ikke lånt fra mnt. ifl. Hellquist og N. Å. Nielsen).

Begriffsdefinition:

Rechtsinstitut: Zwangsvollstreckungsmaßnahme (Person oder Sache).

Metasprache: Neuhochdeutsch

Nhd. Bezeichnung: Subst. Beschlagnahme, Arrest

Mnd. Termini, chronologisch geordnet: besettinge, besâte f., besâtinge, kumber, bekommeringe, bekûmberinge, bekûmmeringe, beslach (14. Jh.), rostêringe. Weniger frequent: wedde, hinder(e) m., n., hechte.

Aldtân. Termini: væd(dj)e, hinder; pant, besetting, hægte, rustering.

Altschwed. Termini: hinder; pant, besate (m.?), besata f., besating, bekymbring, bekymmer, häkte, rostering.

Nhd. Bezeichnung: Verb beschlagnahmen, arrestieren.

Mittellat. Termini: pignorare, vadiare, occupare, detinere.

Mnd. Termini: besetten, besâten, uphalden, kumberen, kumberen, bekummeren, bekummeren, bekommeren (seit dem 13. Jh.), rostêren (15.-16. Jh.), arrestieren (17. Jh.). Weniger frequent in der vorliegenden Bedeutung in den Verwaltungstexten der Periode: wedden, hinderen, panden, hechten. Nur von einer Sache: schütten, beslân.

Altdän. Termini: opholdæ, hindræ, vædhæ; pantæ, besettæ, hægte, rostere.

Altschwed. Termini: upphalda, hindra; panta, besâtja, besâta, bekymbra (15. Jh.), häkta, rostera.

Nb! *arrestieren/Arrest* er lånt fra fransk. *Beschlagnahme* dannedes i tysk i begyndelsen af det 19. årh., verbet *beschlagnahmen* efter substantivet i midten af det 19. årh.

For ovenstående retsbegreb findes der 32 termer på tysk. I oversigten her er kun de termer medtaget der har interesse for sammenligningen af mnt. og glda./fsv. forvaltningsprog.

Denne oversigt svarer nærmest til position 7 hos Reichmann: "Onomasiologisk felt". Hvordan betegnes det aktuelle retsbegreb i de sprog der sammenlignes? Efter min mening er det mest formålstjenligt at beskrive de forskellige termer der hører til samme synonymfelt på eet sted og ikke - som man ellers gør - opføre disse termer på respektive steder i alfabetet med et væld af krydshenvisninger. For hver af de i synonymfeltet optagne termer angives derefter de for de Reichmannske positioner 1-4 relevante data, som tilsammen vil kunne give et billede af den pågældende retsterms geografiske og kronologiske udbredelse, hvilket er af afgørende betydning for konstateringen af forholdet mellem højtyske og middelnedertyske retstermer og de sidstnævntes udbredelsesveje i de tilfælde, hvor de har vundet indpas i de nordiske sprog i middelalderen.

M.h.t. de fonologisk/grafemiske data volder den specielle praxis hos middelnedertyske skrivere at undlade at betegne omlyd ved *o* og *u* vanskeligheder. Endvidere er det sådan at nogle mnt. dialekter har omlyd, andre ikke. Man må altså undersøge, om den aktuelle form kommer fra et dialektområde med omlyd eller ej og desuden, inden man vælger lemmaformen, konstatere om den omlydte form er den mest repræsentative for perioden. Schiller og Lübben havde som jurister ingen anelse om omlydsforholdene i middelnedertysk. De citerer direkte de mnt. skrive-re, der altså gennemgående ikke betegnede omlyd af *o* og *u*, hvorfor Schiller-Lübben som lemma har den uomlydte form. Desværre citerer alle andre ordbøger Schiller-Lübben, når de skal anføre et mnt. ord eller en mnt. form. Det gælder ODS, Kalkar, Niels Åge Nielsens etymologiske ordbog og sågar Gammeldansk ordbogs prøveartikel om *kværn* og *mylnæ* 1992 (S.22-23). Her angiver man under *mylnæ* som mnt. form *mol(l)e*, skønt den korrekte form er *möle* (cf. ht. Mühle), en form man for øvrigt kan finde hos Lasch-Borchling-Cordes som lemma. Marius Kristensen var klar over at der fandtes både omlydte og ikke-omlydte former af samme ord i mnt., men han troede at de ikke-omlydte former af de mnt. låneord i gammeldansk repræsenterede et ældre stadium, så at f. eks. gammeldansk *provæ* var et ældre

lån end gammeldansk *pröva* fra henholdvis lat. *probare* og mnt. *pröven*. I virkeligheden kan det gammeldanske *provæ* lige så godt være et eksempel på de mnt. skrivers praksis at undlade at markere omlyd ved *o*. Af de her nævnte mnt. ord har Lasch-Borchling-Cordes den omlydte form som lemma: *bepröven, bekümmeren*.

De øvrige positioner behandler jeg her kun kort: I den morfologiske position angives de sædvanligvis anvendte forkortelser for køn, bøjning etc., hvis de ikke allerede er angivet i position 1. I positionen for betydningsangivelser må der bl.a. redegøres for, hvorvidt de optagne synonymmer dækker hinanden helt eller kun delvis. F. eks. anvendes de gamle termer mnt. *wedden/wedde, panden/pant* næsten kun om egentlig "pantsætning". Det samme gælder de gl.danske og fsv. parallelle termer.

Ved sammenligningen af de mnt. og de nordiske termer for at "*erklære nogen fredløs*", må man gøre opmærksom på at det semantiske indhold i de identiske termer er helt forskelligt, idet *fredløshedsbegrebet* defineredes helt forskelligt i Tyskland og Norden i middelalderen. "Freden" = lovbeskyttelse (Rechtsschutz) er i Norden først beskyttelse af individet ved slægten, senere ved retssamfundet og oprindeligt altså ikke beskyttelse ved kongen, samfundet eller kirken som i Nordtyskland. Den mnt. term "*vredeløs leggen*" betegner altså noget helt andet end den tilsvarende fsv. term "*læggia/göra friþlösæn*" (gl. da. *frithløs swæræ*).

For øvrigt kan man heller ikke udgå fra at f. eks. identiske mnt. termer har samme betydningsindhold, da deres semantiske værdi retter sig efter hvilket lokalt retssystem (jurisdiktion) de er indplaceret i. Hvis det er aktuelt, vil der blive redegjort for hvilke syntagmer respektive retstermer indgår i samt hvorvidt disse setninger er gået over i det folkelige talesprog i form af faste vendinger og talemåder. Endelig vil der, hvor det er påkrævet, blive givet *etymologiske oplysninger* til de enkelte termer. F. eks. bør mnt. *schütten* i det behandlede ordfelt for "Arrest" forklares, da det etymologisk er en fuldstændig parallel til det fra Rheinland stammende *kumberen*. Mnt. *Schutt* betegner oprindeligt 'Aufstau des Wassers', derefter 'Hinder', 'Arrest'.

Der er altså nok at gøre. Foreløbig er jeg i gang med en snes ordfelter. Hvad der fremfor alt tager tid, er at gennemarbejde de juridiske undersøgelser, der sammenligner tysk og nordisk ret i ældre tid.

Litteratur

Deutsches Rechtswörterbuch. Hg. v. der Preussischen Akademie der Wissenschaften. Weimar 1914 ff., jetzt von der Heidelberger Akademie der Wissenschaften.

Frühneuhochdeutsches Wörterbuch. 1986. Hg. v. Robert R. Anderson, Ulrich Goebel, Oskar Reichmann. Berlin, New York.

Handwörterbuch zur deutschen Rechtsgeschichte. Hg. v. Adalbert Erler und Eckehard Kaufmann. Berlin 1971ff.

Hyldgaard-Jensen, Karl: *Bericht des Projekts EDV-Analyse mittel-niederdeutscher Rechtstexte*. I: Kopenhagener Beiträge zur germanistischen Linguistik 1, S. 7-22. Kopenhagen 1972.

Jørgensen, Merete K., Nielsen, Marita Akhøj, Sørensen, John

- Kousgård: *Gammeldansk Ordbog*. Det danske Sprog- og Litteraturselskab, Præsentationshæfte 21. København 1992.
- Mittelniederdeutsches Handwörterbuch*. Hg. v. Agathe Lasch, Conrad Borchling und Gerhard Cordes, jetzt v. Dieter Möhn. Neumünster 1928 ff.
- Moberg, Lena: *Lågtyskts och svenskt i Stockholms medeltida tänkeböcker*. Uppsala 1989.
- Schiller, K. und Lübben A.: *Mittelniederdeutsches Wörterbuch*.
Bd. 1-6. Bremen 1875-81.
- Skautrup, Peter: *Den jyske Lov*. Text med oversættelse og kommentar. Ordbog S. 210-389. Udg. af Jysk Selskab for Historie, Sprog og Litteratur. København 1941.
- Wörterbuch der mittelhochdeutschen Urkundensprache auf der Grundlage der altdeutschen Originalurkunden bis zum Jahr 1300*. Unter Leitung von Bettina Kirschstein und Ursula Schulze erarbeitet von Sibylle Ohly und Peter Schmitt. Berlin 1986f.

Parsing af ikke-strukturerede ordbogsdata

Claus Bo Jørgensen

Abstract

Firmaet TEXTware har gennem de seneste 4 år konverteret traditionelle data til et stort antal ordbøger til struktureret tekst, dvs. tekst med eksplicit markering af en hierarkisk struktur, i overensstemmelse med den internationale SGML-standard. SGML-teksten er blevet anvendt i redigeringsystemet GestorLEX eller som et led i produktionen af elektroniske ordbøger eller til begge formål.

Erfaringerne fra dette arbejde peger på et antal typiske problemer i forbindelse med parsing, dvs. automatisk strukturering, af eksisterende, ikke-strukturerede ordbogsdata - typisk fotosætter- eller tekstbehandlingsfiler. Problemerne spænder fra banale tegnsætningsfejl over strukturer, der er problematiske for en maskine men ikke for en menneskelig læser, til principielt uløselige fortolkningsspørgsmål.

De enkelte problemtyper og mulige løsninger bliver præsenteret, og det diskuteres, hvordan og i hvilket omfang erfaringerne fra parsing kan udnyttes til at højne kvaliteten af såvel trykte som elektroniske ordbøger.

Indledning

Med ikke-strukturerede ordbogsdata menes i denne artikel f.eks. ordbogsdata skrevet med et almindeligt tekstbehandlingsprogram. Disse data er struktureret i den forstand, at der er mere eller mindre præcist formulerede redaktionsregler, der foreskriver i hvilken rækkefølge de forskellige typer af information skal gives, og hvordan de skal præsenteres (tegnsetning, skrifttype, skriftsnit). De er imidlertid ikke eksplicit struktureret, idet tekstbehandlingsfilen kun indeholder oplysninger om formateringen, dvs. den ønskede behandling, ikke om den logiske struktur.

Data af denne type har vi levet med i flere århundreder, og det betragtes ofte som helt problemfrit at læse og forstå sådanne data. Dette gælder i nogen grad, så længe læseren er et menneske. I det øjeblik man sætter et EDB-program til at fortolke sådanne data, f.eks. for bedst muligt at kunne præsentere dem og søge i dem i en elektronisk ordbog, afsløres der midlertid en hel række af problemer. En del af disse knytter sig selvfølgelig til EDB-programmets begrænsede "evner", mens mange også er reelle problemer for en menneskelig læser. EDB-programmet er blot med sin konsekvens og enfoldighed særlig velegnet til at afsløre dem.

De ideelle ordbogsdata til udnyttelse via et EDB-program indeholder eksplicit mar-

kering af, hvilken type af information de enkelte dele af en artikel giver, og hvordan delene hænger sammen. Den formalisme TEXTware har valgt til strukturbeskrivelsen er SGML, en international standard (se f.eks. Norling-Christensen, 1992), men formalismen i sig selv er ikke vigtig, hvis blot to væsentlige principper er overholdt:

- 1 det der markeres er den logiske struktur: hvilken type af information et stykke tekst giver - ikke hvordan teksten skal præsenteres (eller på anden måde behandles) til et givet formål.
- 2 der er tale om en hierarkisk struktur, en træstruktur, med etiketter klistret på de enkelte mindste enheder i artiklen og med non-terminale elementer, altså knuder der hver samler et antal mindre enheder.

En ordbogsartikel som struktureret tekst vil typisk indeholde en hel del mere information (en finere inddeling) end den tilsvarende traditionelle, typografisk markerede tekst, og skift mellem de større logiske enheder falder ikke nødvendigvis sammen med typografiske skift.

Spørgsmålet er nu, hvordan man automatisk kan komme fra typografisk markeret tekst til struktureret tekst.

Fra formateret til struktureret tekst

Svaret på dette spørgsmål er, at det kan man typisk ikke, hvis man tager ordet automatisk helt bogstaveligt. Der vil altid (efter vores erfaring) være artikler, der ikke umiddelbart passer ind i den struktur, man forsøger at genkende - og der vil være artikler der kan analyseres på flere forskellige måder.

Som basis for selve parsingen har man (i bedste fald) det sæt redaktionsregler der har ligget til grund for ordbogsskrivningen og som foreskriver:

- hvilke oplysninger der gives
- hvilken rækkefølge
- med hvilken skrifttype og -snit
- afgrænset med hvilke skilletegn
- evt. taget fra et lukket vokabular

Opgaven går så ud på at kombinere ovenstående regler med de faktisk foreliggende data, og sætte en parser til herudfra at finde ud af hvad leksikografen har tænkt og ment!

Processen kan deles op i to ikke helt uafhængige dele:

- genkendelse af informationstyper for den enkelte tegnstreng (afgrænsning af terminale elementer - svarende til tag-og-tekst (feltnavnindhold) data)
- sammenhægtning af informationstyper i træstruktur afgrænsning af non-terminale elementer - hvilke informationer hører sammen)

Genkendelse af informationstyper kan i en del tilfælde foretages rent lokalt, det vil sige uden at kigge på den kontekst en tegnstring står i, ud fra selve tegnstringen og dens formatering. Dette vil typisk være tilfældet ved f.eks. stil- og emneangivelser, der tages fra en lukket liste.

Det lader til at være en god idé at foretage så meget som muligt af denne type genkendelse, inden den egentlige strukturparsing starter.

Selve parsingen omfatter derefter genkendelse af de resterende (ikke entydige) informationstyper, og sammenhægtningen af disse i en træstruktur.

Afhængig af formålet med parsingen kan man vælge enten at få så mange som muligt af ordbogens artikler genkendt og struktureret med så få ændringer som muligt, eller at få artiklerne passet ind i en mere konsekvent struktur, med et større antal håndrettelser til følge.

Til parsingen anvendes en grammatik, der kan betragtes som en udvidelse af den SGML-grammatik, der beskriver den resulterende struktur. Udvidelserne er det, der sørger for at genkende kombinationer af skilletegn, formatskift og kontekstinformation som implicitte angivelser af den logiske struktur.

For at give et indtryk af arbejdets omfang kan det oplyses, at en SGML-grammatik på 1 1/2 til 2 A4-sider typisk vil modsvares af en parsegrammatik på 20-25 A4-sider.

Typer af problemer

Under selve parsingen (eller senere) viser problemerne sig på tre måder:

- 1 ikke-analyserede artikler
- 2 flertydige artikler: flere strukturer foreslået
- 3 kun én analyse, men forkert struktur foreslået

Inden for hver gruppe er der igen flere typiske årsager til problemet:

1) Ikke-analyserede artikler:

- fejl i tegnsætning

Tegnsætning er den sværeste del af en ordbog at læse korrektur på, og det er ofte let for et menneske at forstå meningen på trods af en formel fejl.

NB: en del fejl kan normalt fjernes ved at gøre parse-grammatikken mindre restriktiv (dvs. acceptere en afvigende tegnsætning som implicit markering af den samme struktur), men resultatet af en sådan ændring bør kontrolleres grundigt, da man risikerer at introducere fejl af typen entydig men forkert analyse!

- fejl i fontmarkering

Disse er lettere at fange ved korrekturlæsning. Til gengæld kan de meget let opstå når man tilføjer eller sletter data i et tekstbehandlingsprogram.

- uventet struktur

Redaktionsregler er åbne for fortolkning. Dette viser sig tydeligst i større ordbøger med skiftende redaktører eller redaktionsgrupper.

2) Flertydige artikler:

- flere fortolkninger af en enkelt tegnstreng
- Ofte anvendes samme formatering til en del forskellige informationstyper, også med samme position i data. Det kan f.eks. være skellet mellem optionel del af kildesprogsfrase og ikke-tabellagt betydningsspecifikation: begge i kursiv med parentes omkring.
- flere fortolkninger af tegnsætning
- Spørgsmålstegn og udråbstegn som rene indholdstegn (del af kildesprogsfrase eller oversættelsesækvivalent) eller som kombinerede indholds- og skilletegn (i så fald: svarende til adskillelse med komma eller semikolon?).
- Komma som indholdstegn eller som skilletegn.
- flere mulige placeringer (niveauer) i strukturen
- Samme type af information kan ofte placeres mere eller mindre højt i strukturen, dvs. gældende for en større eller mindre del af artiklen. En betydningsspecifikation kan således fortolkes som gældende for den nærmest følgende oversættelse alene, eller for alle de følgende oversættelser.

3) Forkert analyserede artikler:

- fejl i tegnsætning
- fejl i skrifttype eller skriftsnit
- ikke-genkendt flertydighed

Praktiske løsningsforslag

Problemerne med ikke-analyserede eller forkert analyserede artikler kan siges at være principielt uinteressante.

Hvis det er tegnsætningsfejl eller fejl i brugen af skrifttype/-snit der er årsagen, skal disse blot rettes og ariklerne parses igen.

Fejl i form af uventede strukturer kræver et valg mellem rettelse i data og udvidelse af grammatikken. Valget afhænger i meget høj grad af, hvad den strukturmarkerede tekst skal bruges til. Som oftest har opgaven for vores firma været: lav en elektronisk bog - billigst muligt - med alle de eksisterende data i, og det har givet meget lidt restriktive grammatikker. Hvis formålet derimod er at gå over til at redigere ordbogen på struktureret form, vil man ofte benytte lejligheden til at rydde lidt op i de eksisterende variationsmuligheder.

De interessante problemer er de flertydige konstruktioner. Også på dette område bærer TEXTware's løsningsforslag præg af, at ekstra arbejde (højere pris) skal kunne forsvares med en tilsvarende gevinst i udnyttelsen af data.

Desuden kræver de principielt flertydige konstruktioner en egentlig leksikografisk stillingtagen, og det er ikke vores job.

Der er derfor i praksis anvendt følgende "løsnings"-modeller:

Flertydig informationstype (f.eks. betydningsspecifikation vs. optionel del af kilde-sprogsfrase):

Eksempel:

se sig godt for (før man handler)

[y] regarder à deux fois

Frase *se sig godt for*
(før man handler)

Frase *se sig godt for*

Tyd1

Tyd1

TydSpec (*før man handler*)

Tyd2 [y] regarder à
deux fois

Tyd2 [y] regarder à
deux fois

I en passiv ordbog vil de to muligheder anvende hver sit sprog, og i praksis har det vist sig at en ret kort liste med hyppige ord (præpositioner, artikler, evt. pronominer og enkelte substantiver) er nok til at kunne afgøre, hvilket sprog der er anvendt.

Denne flertydighed har vi derfor normalt taget stilling til i hvert enkelt tilfælde.

I en aktiv ordbog er der ikke mulighed for at skelne mellem funktionerne ved hjælp af sproget. Den eneste mulighed ligger i en egentlig "forståelse" af teksten.

Da vi i de elektroniske ordbøger plejer at gøre både kildeprog og metasprog søgbare, er der ikke noget umiddelbart tab af funktionalitet ved ikke at skelne mellem disse, så en flertydighed hér er blevet opløst ved altid at vælge den ene fortolkningsmulighed (del af kilde-sprogsfrase).

Flertydig tegnsætning:

Eksempel:

Voulez-vous de la bière? -- Va pour la bière! Vil De have øl? -- ja, lad gå!

allez toujours! bliv bare ved! hæng bare i! lad Dem ikke forstyrre!

est-ce qu'on va au cinéma, dis, papa? skal vi så i biografen, hvad, far?

on dirait man skulle tro, man kunne fristes til at sige;

on dirait de loin une barque på afstand kunne man godt tro, at det er en båd; det kunne godt ligne en båd på afstand

allons donc! hvor vil de hen? hvad behager! nej hør nu! (*iron.*) javist! nå! nå skynd dig nu! (*beroligende*) nå nå!

allez! (sp.) start! (*i boksning*) time!

Ved spørgsmålstegn og udråbstegn er der, afhængig af hyppigheden, enten taget stilling

til hvert enkelt tilfælde, eller der er konsekvent valgt én af mulighederne (betydningsadskillelse som ved komma eller som ved semikolon - skellet mellem rent indholdstegn og kombineret skille- og indholdstegn vil ofte fremgå af konteksten).

Komma har vi normalt betragtet som skilletegn, men igen er der ikke nogen forskel med hensyn til funktionalitet i de elektroniske ordbøger, idet vi plejer at bruge semikolongrænsen som kriterium for linieombrydning. Det kan altså heller ikke ses i den elektroniske bog, om kommaet er skille- eller indholdstegn.

Flertydigt niveau i strukturen:

Eksempel:

I. **navle** *en -r*

1 (*anat*) navel; (*fagl.*) umbilicus; T (*lettere vulg.*) belly-button;

2 (*bot.*) hilum

Krop

Tyd1

| TydSpec *anat*

|

| Tyd2

| Ækvi navel

| Tyd2

| TydSpec *fagl.*

| Ækvi umbilicus

| Tyd2

| TydSpec *lettere vulg.*

| Ækvi belly-button

Krop

Tyd1

Tyd2

| TydSpec *anat*

| Ækvi navel

Tyd2

TydSpec *fagl.*

Ækvi umbilicus

Tyd2

TydSpec *lettere vulg.*

Ækvi belly-button

Endnu en gang har de praktiske hensyn fået overtaget: der sker mindst skade ved en forkert fortolkning, hvis man vælger den højest mulige placering (det størst mulige virkefelt). I værste fald giver dette fejl, som er gennemskuelige for den menneskelige bruger (to modstridende TydSpecifikationer knyttet til samme oversættelsesækvivalent), mens valg af det mindst mulige virkefelt giver "usynlige" fejl i form af manglende information om de følgende ækvivalenter (i eksemplet ovenfor vil 'umbilicus' ikke kunne findes som anatomisk term).

I forbindelse med denne sidste type af flertydigheder kan det bemærkes (kommentar fra Ole Norling-Christensen), at endog den leksikograf, der har skrevet artiklen, kan være i tvivl om den rette fortolkning. Dette blot for at minde om i hvor høj grad de traditionelle værktøjer til ordbogsredigering har fokuseret på *præsentation* snarere end på *struktur*.

Problemerkernes hyppighed

Et lille antal stikprøver fra forskellige af TEXTwares parse-opgaver afslører, at de fleste problemer, der kræver manuel indgriben i processen (rettelse i data eller markering af én af de mulige analyser som den foretrukne), skyldes de "banale", dvs. principielt uinteressante, fejltyper. Typisk står tegnsætningsfejl for 30-40% af tilfældene, fejl i skrifttype/-snit for yderligere 20-30%, mens flertydigheder og egentlige strukturfejl (brud på redaktionsreglerne) hver giver 10-20%. Endelig er der ofte en mindre gruppe af egentlige kodningsfejl, dvs. fejlkodning af specialtegn eller lignende.

Fordelingen på de enkelte problemtyper skyldes i høj grad kontraktmæssig afgrænsning af TEXTwares ydelser over for forlagskunderne. Den "interessante" flertydighed med hensyn til en betydningsspecifikations niveauplacing er således typisk langt det hyppigste problem, men en fuldt tilfredsstillende løsning kræver som nævnt egentlig leksikografisk stillingtagen.

Konklusion

En automatisk behandling af traditionelle, formaterede ordbogsdata afslører altså en række problemer ved fortolkningen af disse. De banale problemer som fejl i tegnsætning og valg af skrifttype/-snit kan let løses, men understreger dog et rimeligt krav til ethvert ordbogsredigeringsværktøj: præsentationen af artiklerne skal ikke være en del af selve ordbogens data men alene afspejle den logiske struktur.

De interessante problemer, flertydighederne, kræver egentlig leksikografisk stillingtagen og er derfor dyre at løse. Værdien af en teoretisk tilfredsstillende løsning afhænger desuden i høj grad af den planlagte udnyttelse af de strukturerede data.

Strukturoplysningerne er afgørende, når data skal behandles automatisk, som f.eks. ved udnyttelse i forbindelse med maskinoversættelse (meget kritisk) eller blot i elektroniske bøger, hvor maskinens rolle er at formidle ordbogens data til en menneskelig læser.

Desuden må det konkluderes, at muligheden for at udnytte strukturen under præsentationen af data for en menneskelig bruger, er større ved elektronisk formidling end ved udgivelse i trykt form. For det første er der de traditionelle hensyn at tage til læseligheden (et roligt skriftbillede uden for mange forskellige skrifttyper og skriftsnit) og den trykte bogs omfang (omend tendensen går i retning af mindre komprimeret, mere "luftig" opsætning). For det andet er der den grundlæggende forskel, at den én gang trykte præsentation ikke kan ændres af læseren, hvis en anden synsvinkel på data ønskes! I den elektronisk formidlede form koster en luftig opsætning derimod ikke noget, og læseren kan selv (i nogen grad) vælge den præsentation af data der passer til det aktuelle formål.

Det er således ikke tilfældet, at en trykt udgave af en ordbog nogensinde vil kunne udnytte den strukturelle information i ordbogsdata i samme grad som en elektronisk udgave. Det er imidlertid stadigvæk en glimrende idé at det er de samme strukturerede data der ligger til grund for begge udgivelser.

Litteratur

Ole Norling-Christensen, 1992. *Struktureret redigering af ordbøger*, i: Nordiske studier i Leksikografi. Nordisk forening for leksikografi, Oslo, 1992.

Den Danske Ordbogs tekstkorpus og spORDhunde

Kjeld Kristensen

En spORDhund er et væsen, der har snuden i niveau med sprogets græsrodde eller lidt over, og med vibrerende næse og ører er på jagt efter ord, altsammen med henblik på det overordnede mål, Den Danske Ordbog. Den Danske Ordbog har kontakt med 600 spORDhunde, personer, der registrerer nye måder at bruge sproget på og indsender resultaterne til os på redaktionen. Jeg vender tilbage til spORDhundene, men til dem har vi netop udsendt det nyhedsbrev, I har fået et eksemplar af. På bagsiden kan man læse en punktvis præsentation af Den Danske Ordbog, og den kan måske være nyttig at se på som optakt til mit og de to følgende foredrag.

Den Danske Ordbog

- er en ordbog over dansk sprog i perioden fra ca. 1950 til i dag, med hovedvægten på de sidste ti år
- skal rumme over 100.000 opslagsord
- skal indeholde oplysninger om ordenes stavning, ordklasse, bøjning, udtale, betydning, konstruktion (forbindelser med andre ord) og etymologi (ordenes oprindelse og historie), og autentiske citater skal belyse ordenes brug
- skal vise sproget, som det er, men også vejlede brugeren
- skal være et udbredt og populært værk, tilgængeligt for både den professionelle og den utrænede bruger
- udarbejdes på grundlag af en elektronisk tekstsamling på 40 millioner løbende ord, andre ordbøger og leksika, Dansk Sprognævns seddelkartotek og mange tusinde ordsedler indsendt af flere hundrede spORDhunde
- redigeres af de tre ledende redaktører mag.art. Ebba Hjorth, cand.mag. Kjeld Kristensen og mag.scient. Ole Norling-Christensen samt en halv snes redaktører og konsulenter
- udgives af Det danske Sprog- og Litteraturselskab og finansieres af Kulturministeriet og Carlsbergfondet i fællesskab
- vil koste ca. 2000 kr. og udkommer i 6 bind på forlaget Gyldendal i 1998-1999

Mit indlæg skal især handle om ordbogens elektroniske tekstkorpus. Et elektronisk tekstkorpus kan defineres som en samling elektronisk læsbare tekster, sammensat ud fra nogle bestemte kriterier med henblik på et bestemt formål. Formålet er i dette tilfælde givet ved, at korpus skal gøre det muligt for os at gennemføre de lingvistiske undersøgelser og analyser, der sætter os i stand til at løse den opgave, vi har fået stillet, nemlig at redigere ordbogen. At korpus så også kan og bør bruges til mange, mange sprogvidenskabelige, litteraturvidenskabelige og andre undersøgelser, er så en anden sag. Men en konsekvens af det formål, som korpus skal sammensættes efter, er, at korpus på en eller anden måde skal afspejle moderne dansk sprog i perioden 1983-92. Korpus skal være repræsentativt for det tekstunivers, der hedder dansk 83-92 - eller skal det?

For det første er repræsentativitet, i hvert fald kvantitativ repræsentativitet, et idealt krav. For det andet vil jeg hævde, at vi ved opbygningen af vores korpus hverken kan eller skal stræbe efter at opfylde dette ideale krav. Man kan nemlig stille sig selv en række spørgsmål, fx: Tales der flest ord eller skrives der flest ord en tilfældig dag i Danmark? Der tales langt flere ord, end der skrives. Skal der så ikke være mest talesprog i vores korpus (forudsat vi også ønsker at beskrive talesproget i ordbogen - hvad vi gør)? Nej, dels vil det naturligvis være alt, alt for ressourcekrævende at optage og udskrive alt det talesprog (især da langt det meste produceres face-to-face i hverdagens almindelige situationer), dels ville man derved få underbelyst skriftsproget, der jo oftest udbredes videre end talesproget; det er især de professionelle skribenters, bl.a. journalisters og forfatteres kommunikation med de mange i aviser, blade og bøger. Og hvordan skal vi repræsentere professionelt sprog i henseende til medium, genre og emne: skal vi vægte efter oplagstal og seer- og lyttertal? Nej, repræsentativitet er umulig at opnå, og den er heller ikke ønskelig. Det, det gælder om, er at etablere en afbalanceret samling eksempler på dansk sprog brugt i tale og skrift, som almensprog og som fagligt sprog, spredt over et stort antal medier og genrer og emner, og produceret af danskere af begge køn, i forskellige aldre og med forskellig uddannelse og erhverv. Det skal, så vidt ressourcerne rækker, være et eksemplarisk, afbalanceret korpus. Det er bredden, alsidigheden, der tæller. Og skulle man så ved brugen af korpus opdage, at sammensætningen er lidt skæv på et punkt eller to, kan man jo løbende udskifte nogle af teksterne med andre - eller supplere.

På to punkter har vi bestræbt os på at give vores korpus en særlig drejning, som ligger i forlængelse af den bredde, der skal præge ordbogen. Den skal være bred i anlægget, dvs. vise sproget, som det er (en deskriptiv ordbog), og ordbogen skal være et udbredt og populært værk. Derfor skal det ikke bare være det professionelle skriftsprog, der beskrives i ordbogen; Den Danske Ordbog skal ikke kun være en litterært baseret ordbog som Ordbog over det danske Sprog, domineret af citater fra tidens førende penne. Vi bygger så vidt muligt også på almindelige menneskers dagligdags sprog, ikke mindst deres talesprog. Vi lægger vægt på at få meget produktion med i vores korpus, altså sprog, som folk selv frembringer, i modsætning til det sprog, som folk læser eller hører, (reception). Og vi lægger vægt på at få meget talesprog med. Vi samler breve, dagbøger, lejlighedssange, læserbreve, opgaver og danske stile som eksempler på produktion, og vi

samler interviews og gruppesamtaler fra adskillige sprogvidenskabelige og andre projekter. Desuden en mængde udskrifter af talesprog i radio og tv og af taler i Folketinget og Københavns Borgerrepræsentation. Vi etablerer en samling talesprogstekster, der siger sparto til det meste af, hvad der findes af den slags i hele verden: ca. 5 mio. løbende ord. Alligevel vil professionelle skribenters sprog blive dominerende i korpus, deres tekster er lettest og billigst at få fat på, men alligevel...

Det praktiske arbejde med at indsamle tekster og opbygge korpus foregår ved, at vi tager kontakt med avisredaktioner, bladredaktioner, forlag og andre instanser, vi regner med har noget at tilbyde. Vi har overalt mødt en kolossal imødekommethed og velvilje uden økonomiske bagtanker. Vi får så teksterne på diskette eller tape, i et eller andet format, konverterer fra dette format til DOS, retter teksterne til i WordPerfect efter vores standarder og lægger dem ind i vores database, PARADOX. Hver tekst forsynes med en såkaldt header med oplysninger om teksten. Jeg vil kort gennemgå, hvilke oplysninger sådan en header indeholder.

Header

Kildeangivelse

Tekstgruppe

entydig identifikation af en gruppe (beslægtede) tekster

Tekstnr

løbenummer inden for tekstgruppen

Restriktion?

Restriktion_a

tekstens navne skal anonymiseres: "ja"/"nej"

Restriktion_b

teksten må ikke bruges til andet end ordbogen: ja"/"nej"

Udløbsår

for restriktion b

Sprogbruger+

Rolle?

især ved talesprog, fx "interviewer og "interviewperson"

Identifikation?

entydig tretegnskode, hvortil replikker i talesprogstekster kan referere

Efternavn?

hvis det kendes

Fornavn?

hvis det kendes

*Køn

"m"/"k"/"u[kendt]"

Uddannelse?

hvis den kendes

Erhverv?

hvis det kendes

*Fødselsår?

et heltal mellem 1880 og 1990, hvis året kendes

Sikker?

"?", hvis året ikke kendes med sikkerhed

Fødested?

hvis det kendes

*Sprogvariant

"rigssprog"/"regionalsprog"

Teksttitel?

hvis der er en titel

Værktitel?

navn på fx antologi, avis, blad, hvis relevant

Forlag?

normalt kun ved bøger og radio-/tv-stationer

| | |
|-------------------------|---|
| Datering | |
| Dag? | hvis den kendes |
| Måned? | hvis den kendes |
| *År | et heltal mellem 1983 og 1992 (begge inkl.) |
| Sikker? | "?", hvis året ikke kendes med sikkerhed |
| Lokalisering? | fx sektion/side/spalte i avis; (bind og) side ved bøger |
| Tekstbeskrivelse | |
| *Sprogtype | "almensprog"/"(alment) fagsprog" |
| *Udtryksmedium | "skrift", "tale" eller et par mellemformer |
| *Synsvinkel | "reception"/"produktion" |
| *Aldersrelation? | "barn-barn"/"barn-ung", "barn-voksen"/.. /"voksen-voksen" |
| *Medium | ét fra en liste med 12 medier, bl.a. bog, avis, blad, radio |
| *Genre? | én fra en liste med 124 delvis medieafhængige genrer, fx roman, brev, interview, nyhedsudsendelse |
| *Emne? | ét fra en liste med 64 emner, fx biologi, litteratur, politik, fysik |
| Omfang | antal løbende ord i teksten (beregnes maskinelt) |

Indrykningerne viser den hierarkiske struktur. Kildeangivelsen indeholder nogle facts om teksten, som umiddelbart foreligger, eller som kan slås op i egnede håndbøger. Tekstbeskrivelsen er mere en ekstern sproglig beskrivelse af teksten mht. en række parametre. * betyder, at det, man skal skrive i det pågældende felt, er hentet fra et lukket univers; man skal altså vælge én ud af en begrænset mængde bestemte værdier. + angiver, at feltet er multipelt (der kan være flere sprogbrugere), ? at feltet er fakultativt.

Meningen med headeroplysningerne er, at man maskinelt opbevarer de oplysninger om teksten, man senere får brug for, fx som kildeoplysninger ved citater, eller som man skal bruge ved de lingvistiske analyser af ordenes brug og betydning, fx en undersøgelse af, om et ord i den og den betydning især bruges af børn og unge, især i talesprog, osv. Man kan også foretage selektive søgninger i korpus, fx af mænds sprogbrug i læserbreve, eller af det faglige sprog i tekster om biologi. Det er bl.a. her, man har glæde af felterne med det lukkede univers af værdier.

Til sidst lidt om spORDhunde-initiativet. Det har et engelsk forbillede i Word Watcher-kampagnen. Men vi har altså valgt et sjovere og mere kreativt ord. Vi havde en artikel i brugsforeningsbevægelsens blad "Samvirke" i september 92, med en præsentation af ordbogen og en opfordring til læserne om at henvende sig til os, hvis de var interesserede i at finde ord til os. Der kom 600 henvendelser pr. telefon svarer og brev. Vi udsendte informationsmateriale og ordsedler til alle, og der er indtil nu kommet 3-4000 udfyldte ordsedler retur fra over 200 spORDhunde. Hensigten med kampagnen er at få knyttet

Kjeld Kristensen

en masse sproginteresserede danskere til Den Danske Ordbog - det er jo fremtidige kunder, men det viser sig også, at det materiale, vi får ind, er virkelig værdifuldt. Vi har i nyhedsbrevet skrevet en lille artikel med de første resultater: nye ord, etablerede ord og udtryk med ny betydning, nye slangudtryk. Det er kun et lille og lidt tilfældigt udvalg; der er som sagt kommet flere tusinde sedler ind. Oplysningerne fra sedlerne lægges ind i en database til senere brug. Og kampagnen fortsætter.

Grammatiske informationer i ordbøger

Hans-Peder Kromann

1. Spørgsmål

Hvordan skal vi *integrere* grammatik i ordbøger? Hvordan skal vi *præsentere* grammatiske informationer i ordbøger? I hvilke ordbogstyper kunne vi tænke os *selvstændige grammatikafsnit*? Det er de tre spørgsmål, som interesserer mig her.

Med grammatik i ordbøger mener jeg først og fremmest syntaks og dertil hørende semantik og sigter ikke til bøjningstabeller for de enkelte ordklasser. Jeg vil anskueliggøre problemstillingen på nogle af de svære tyske præpositioner¹; det er *an, auf, hinter, in, neben, über, unter, vor, zwischen*, der som bekendt styrer enten akkusativ eller dativ.

Præpositioner er et oplagt fælles interesseområde for både leksikografi og grammatikografi. Hverken grammatikere eller leksikografer kan komme uden om præpositioners syntaks og semantik. I grammatikker behandles de i hvert fald som selvstændig ordklasse, og i ordbøger optræder præpositioner med selvstændige ordbogsartikler, hvortil kommer, at de indgår i utallige konstruktioner med verber, adjektiver og substantiver. Ordbøger skal være fulde af præpositioner, og hvis ikke de er det, så er det en mangel ved ordbogen. I det følgende begrænser jeg mig til præpositioner i tyske modersmålordbøger, men problemerne og løsningerne gælder *mutatis mutandis* også for andre sprogs modersmålordbøger samt for tosprogsordbøger.

2. Leksikografi og grammatikografi som komplementære størrelser

I de senere år er der udkommet en del ordbøger med selvstændige grammatikafsnit². Imidlertid er informationerne i grammatikafsnittene kun undtagelsesvis søgt integreret i ordbogsartiklerne. WAHRIG-ordbøgerne har dog et specielt design, hvor der ved hjælp af indekstal etableres krydsreferencer mellem ordbogsartikler og grammatikafsnit, men ellers er der ikke nogen krydsreferencer mellem disse ordbøgers grammatikafsnit og deres ordbogsartikler. Det er grunden til, at ordbøgers selvstændige grammatikafsnit snarere må anskues som en slags mini-grammatikker, der er trykt i samme bog som ordbogen. Det er sikkert nyttigt for mange brugere at have både en ordbog og en elementær grammatik i en og samme bog, men det kan ikke betegnes som den optimale løsning på at tilføre ordbøgerne flere grammatiske informationer.

Det er en udbredt opfattelse at anskue et sprogs grammatik og dets ordforråd som *komplementære størrelser*. I *grammatik-bøger* forventer man at få beskrevet ordklasserne

og de dertil hørende bøjningsparadigmer, ordstilling, transitivitet og valens, tempus, modus, diatese, o.l. Det er den systematiske beskrivelse, hvor grammatikograferne benytter et kerneordforråd til at anskueliggøre det grammatiske system og hvor de derfor præsenterer ordlister, som ikke behøver at være udtømmende. Det er kun ved de rene grammatiske ordklasser som pronominer, præpositioner, konjunktioner og adverbier, man kan forvente en udtømmende behandling af de til ordklassen hørende ord.

I *ord-bøgerne* derimod forventer man at få flest mulige ords betydninger beskrevet. Hvadenten det drejer sig om alfabetiske eller systematiske ordbøger, så går leksikograferne ud fra ordene. Modersmåleksikografer skal først og fremmest præsentere ordenes betydninger, og fremmedsprogsleksikografer præsentere ækvivalenter. Når leksikograferne endvidere forsyner ordbetydningerne med stilistiske, ortografiske, udtalemæssige samt grammatiske informationer, taler man om universalordbøger³.

Når man i fugleperspektiv kigger på ordbøger og grammatik-bøger fra det sidste halve århundrede, så ser de ud til at være gået hver deres vej. Det er som om den lingvistiske strukturalisme har medvirket til at fjerne leksikografi og grammatikografi fra hinanden. Er der med strukturalismen gået et helhedssyn på sprog tabt? I Tyskland i forrige århundrede turde Jacob Grimm, Daniel Sanders og Hermann Paul være penneførende for store ordbøger og store grammatikker. I Danmark i dette århundrede er det vistnok kun Paul Diderichsen og Aage Hansen, der har turdet prøve begge discipliner.

3. Integreret grammatik: Eksempler og syntaktiske konstruktioner

Ordbøgers grammatiske informationer er efterhånden blevet stabiliseret mht oplysninger om ordklasse og bøjning og som regel også om verbers transitivitet. Dertil kommer implicitte syntaktiske oplysninger under de enkelte opslagsord, i og med der anføres eksempler på ordenes brug.

Velvalgte eksempler er nyttige, fordi de kan illustrere et eller flere fænomener. Eksempler kan fx informere om hjælpeverber, transitivitet, valens, passivbrug, præpositioner og deres rektion. Alligevel er eksempler i sig selv ikke tilstrækkelige. Ordbogsbrugerne kan ikke uden videre generalisere ud fra enkelte illustrerende eksempler til en generel syntaktisk konstruktion. Eksemplerne bør derfor suppleres med generaliserede konstruktionsangivelser. Hvordan det kan gøres, vil jeg antyde med et eksempel på valens-, præpositions- og rektionsbrug fra WAHRIG-dtv (1978). I ordbogsartiklen

aufbauen 1 ... 2 (800) *auf einer Sache aufbauen sich auf eine Sache stützen, auf einer Sache gründen*; das Drama baut auf den Regeln der antiken Dramentechnik auf. 3 ...

gives der vitale oplysninger om verbets valens og ved præpositionen *auf* om den i dette tilfælde påkrævede rektion: De valensbundne sætningsled beskrives ved hjælp af talindekset (800)⁴, som indgår i et indeks over valensmønstre (Satzmuster) foran i ordbogen. Det er gennem konstruktionsangivelsen *auf einer Sache aufbauen*, at der generaliserende demonstreres, hvilken præposition *aufbauen* kombineres med. Endelig viser konstruktionsangivelsen og eksemplet ... *auf den Regeln* ..., hvilken kasus *auf* kræver i den-

ne valenskonstruktion. Såvel valens som rektion er integreret i konstruktionsangivelsen, og samtidig anskueliggøres begge kategorier i eksemplerne. Hvis konstruktionsangivelsen *auf einer Sache aufbauen* ikke var blevet medtaget, ville betydningsforklaringen *sich auf eine Sache stützen* have bidraget til at skabe forvirring om *auf*'s rektion.

Denne form for integration af syntaktiske informationer har Wahrig⁵ indført i tysk leksikografi. Heroverfor står DUDEN-ordbøger, som blot har konstruktionsangivelser, men ikke angivelse af hverken valensmønstre eller transitiv/intransitiv.

4. En ordbogsartikel med præpositionen *über*

I store universalordbøger (1-sprogede såvel som 2-sprogede) er artiklerne med ordene fra de 'rene' grammatiske ordklasser (pronomen, artikel, præposition, konjunktion, adverbier og hjælpeverber, modalverber) som regel lange ordbogsartikler. I mange ordbøger beskrives disse ord imidlertid temmelig rodet og ustruktureret. Potentielle brugere kan næppe forstå eller gennemskue forklaringerne, heller ikke selvom de prøver at forstå ved at analysere eksemplerne. Forklaringerne til de grammatiske ord er ikke særlig stringente eller konsistente hverken med sig selv eller med almindelige 'grammatiske' forklaringer. De kan indeholde modsigelser og utilstrækkelige forklaringer. Undertiden kan det se ud som om præpositioner bliver behandlet på samme måde som substantiver, og man gør deres hovedbetydninger til hovedinddelingskriteriet. Og dermed ser man bort fra disse ords syntaktisk-semantiske egenskaber, som måske med større held kunne lægges til grund for sådanne ordbogsartiklers struktur.

4.1. Duden Universalwörterbuch (1989)

For at give en smagsprøve på en ikke fyldestgørende beskrivelse af præpositioner skal vi se på den tyske præposition *über* i den førende tyske modersmålordbog⁶. I DUDEN UNIVERSALWÖRTERBUCH (1989) har artiklen for *über* en mediostruktur, som primært bygger på en beskrivelse af præpositionens hovedbetydninger og ikke på de grammatiske kategorier, som er kendetegnende for præpositioner. Eksempelvis beskrives betydningerne til I, 3 vha 'Reihen- od. Rangfolge', 'Rangordnung', 'Wert, der überschritten wird', 'Abhängigkeitsverhältnis'. Når man så analyserer ordbogens eksempler nærmere, viser det sig i flere tilfælde⁷ mere at være verbernes (jvf. *verfügen* og *herrschen* i I 3 d) end præpositionens betydning, der er blevet fremanalyseret, hvorimod væsentlige grammatiske egenskaber slet ikke fremdrages. Her gengives det grammatiske skelet samt eksempler fra artiklen for *über* i DUDEN UNIVERSALWÖRTERBUCH.

I. Präp. mit Dativ u. Akk.

1. (räumlich)

- a) (mit Dat.) sie wohnt über uns
- b) (mit Akk.) das Bild über das Sofa hängen

- c) (mit Dat.) der Nebel liegt über der Wiese
- d) (mit Akk.) er legte die Jacke über den Stuhl
- e) (mit Akk.) sie entkamen über die Grenze
- f) (mit Akk.) Tränen liefen ihr über die Wangen
- g) (mit Dat.) sie wohnen über der Straße
- h) (mit Akk.) der Rock reicht über das Knie
- i) (mit Akk.) der Weg führt über die S. hinaus
- j) (mit Akk.) der Zug fährt über Mannheim

2. (zeitlich)

- a) (mit Akk.) ich will über das Wochenende segeln
- b) (mit Akk.) fragen sie mich über acht Tage
- c) (mit Dat.) sie ist über der Arbeit eingeschlafen
- d) (mit Akk.) es ist zwei Stunden über die Zeit

- 3.
- a) (mit Dat.) er liegt über dem Durchschnitt
 - b) (mit Dat.) eine Temperatur über Null
 - c) (mit Akk.) Musik geht ihm über alles
 - d) (mit Akk.) über etwas verfügen, über jmdn herrschen

4. (mit Akk.) Fehler über Fehler

5. (mit Dat.) die Kinder sind über dem Lärm aufgewacht

6. (mit Akk.) er wurde über Gebühr gelobt

7. (mit Akk.) über etwas sprechen, reden

8. (mit Akk.) eine Rechnung über 50 DM

8. (mit Akk.) einen Aufruf über alle Sender bringen

10. (mit Akk.) Fluch über die Morder!

11. (mit Akk.) über etwas weinen, lachen,...

12. (mit Akk.) Kinder über 10 Jahre

II. Adverbium

- 1. a) über einen Meter breit
b) über und über
- 2. Gewehr über!
- 3. den ganzen Tag über fleißig lernen

III. Adjektiv

- 1. vier Mark sind über
- 2. a) kräftemäßig ist er mir über
b) es ist mir über, ihn darum zu bitten.

Ud fra et brugeraspekt er det svært at forestille sig indenlandske eller udenlandske

brugere, der kan få glæde af de grammatiske informationer i en sådan ordbogsartikel. Præpositionens hovedbetydninger (noget med sted, noget med tid og så et antal overførte betydninger) fremgår måske nok af artiklen, men mediostrukturen med dens manglende rektionsregler samt dens *ad hoc*-betydningsforklaringer og dertil hørende eksempler viser, at der ikke er lagt nogen som helst præpositionsgrammatisk analyse til grund for ordbogsartiklen⁸.

4.2. En præpositionsartikel med integreret grammatik

Hvilke grammatiske egenskaber er kendetegnende for præpositioners syntaks og semantik? Præpositionerne tilhører den særlige ordklasse, hvis elementer etablerer syntaktiske og semantiske relationer mellem to eller flere sætningsled. Det er derfor vigtigt at få beskrevet disse relationer mellem præpositionen og andre sætningsled. Naturligvis skal præpositioners hovedbetydninger fortsat beskrives i ordbogsartiklerne, men ikke på samme måde som substantiver, adjektiver og verber. I grammatiktunge ordbogsartikler må der tilføjes relevante syntaktisk-semantiske informationer. Omsat til præpositioner indebærer det, at tre fundamentalt forskellige syntaktisk-semantiske relationer *so* *oder* *so* skal lægges til grund for den grammatiske beskrivelse af dem:

1. Præpositioner manifesterer en **valensrelation** ved
 - 1.1. **præpositionsobjekter**
 - 1.2. **valensbundne adverbialer**
2. Præpositioner i ikke-valensbundne præpositionsgrupper manifesterer en **fri relation** i forhold til det ved valensmønstret betegnede sagforhold. Præpositionsgruppen kan betegne
 - 2.1. **sted**
 - 2.2. **tid**
 - 2.3. **måde, middel, grund, anledning, følge**
3. Præpositioner indgår i **semantiske relationer med andre aktanter** således, at den ved et andet sætningsled betegnede genstand og den ved regimen betegnede genstand indgår i⁹
 - 3.1. en **intralokal** relation (ved hvilevb)
 - 3.2. en **translokal** relation (ved bevægelsesvb)
 - 3.3. en **ikke-lokal** relation (ikke-konkrete vb)

Ved hjælp af disse tre syntaktisk-semantiske relationer kan der udarbejdes mere velstrukturerede ordbogsartikler for tyske præpositioner med dobbeltrektion. Præpositioners hovedbetydninger skal naturligvis fortsat stå på et fremtrædende sted, fx i et 'hoved', og de grammatiske egenskaber skal så gøres transparente for brugerne i en 'hale'¹⁰.

Über

- A Hovedbetydninger anføres og/eller illustreres i et 'hoved': ...
- B 'Hale'
1. **Valensbundet sætningsled**
 - 1.1. **Valensbundet adverbial** ved verber, der betegner **bevægelse**: *über + akk.* über die Brücke fahren, über den Stein stolpern, einen Teppich über jmdn decken, eine Jacke über den Stuhl legen; der Fluß tritt über die Ufer
 - 1.2. **Valensbundet adverbial** ved verber, der betegner **hvile**: *über + dat* das Bild hängt über der Tür; es lag eine Decke über dem Tisch; Sie wohnen über der Straße. Sie trägt einen Mantel über dem Kleid
 - 1.3. **Præpositionsobjekt**: *über + akk.* über jmdn, etw. fluchen, sich freuen, wachen, sprechen, ...
 2. **Frit adverbial**
 - 2.1. Adverbialet betegner **stedet**, hvor handlingen sker: *über + dat.* es regnete über den Gärten; er hatte die Hände über der Brust gefaltet
 - 2.2. Adverbialet betegner **tiden**, hvor handlingen finder sted
 - 2.2.1. *über + dat.* (= während) über der Arbeit schief er ein; seine Mutter ist über seiner langen Reise gestorben
 - 2.2.2. *über + akk.* (= in+D) über acht Tage kehrt er wieder zurück (= während) über das Wochenende segelt er
 - 2.3. Adverbialet betegner **årsag** *über + dat.* (= infolge) über dem Lärm aufwachen; über dem Streit ging die Freundschaft in die Brüche;
 - 2.4. Adverbialet betegner **middel** (= via) *über + akk.* er bekam den Anruf über einen Freund; der Zug fährt über Mannheim; einen Anruf über alle Sender bringen;
 3. **Fraseologismer**
über + akk. über Gebühr; Fehler über Fehler; über Berg und Tal; über die (alle) Maßen; über alles; über kurz oder lang; über und über;
 4. Øvrige funktioner

Dette udkast til 'halen' af en præpositionsartikel viser, hvordan man kan *integrere grammatik* i grammatiktunge ordbogsartikler. Med udgangspunkt i en artikelstruktur med 'hoved' og 'hale' kan der i halen etableres et informativt og instruktivt samspil mellem de udvalgte eksempler og beskrivelsen af de syntaktisk-semantiske konstruktioner, de illustrerer. Med en sådan ordbogsartikel *præsenteres grammatiske informationer* hensigtsmæssigt og brugervenligt, og de står under et opslagsord, hvor brugerne kan forventes at søge¹¹.

Ved grammatiktunge ordbogsartikler kan det endvidere være praktisk at henvise til et selvstændigt grammatikafsnit i ordbogen, hvor brugeren så kan få en nærmere forklaring til (glemte) grammatiske kategorier. Ved tyske præpositioner kunne der således henvises til et afsnit, hvor akkusativ-dativrektion omtales og de anvendte syntaktisk-semantiske begreber bliver forklaret. Leksikograferne skal (bør) jo alligevel med henblik på udarbejdelsen af ordbogen lave en grammatik, som udgør grundlaget for de grammatiske informationer i ordbogen, så hvorfor ikke også i selve ordbogen præsentere udvalgte dele af den grammatik, der lægges til grund for beskrivelsen af de grammatiske ord og de øvrige grammatiske informationer.

5. Grammatiske informationer i universalordbøger

Integration af grammatiske informationer i en ordbog kan ske på mange måder og mange steder i en ordbog, og der vil være forskellige opfattelser af, hvad der skal gælde som ordbogsrelevant grammatik. Det er ren retorik at spørge, om universalordbøger for modersmål og fremmedsprog forventes at indeholde fyldestgørende og opdaterede grammatiske informationer, men det er reelt at spørge, om de kan forventes at indeholde selvstændige grammatikafsnit.

Der forekommer allerede i praksis universalordbøger med et selvstændigt grammatikafsnit, som eksisterer uafhængigt af ordbogen som ordbog, og som slet ikke bruges til noget som helst i ordbogsartiklerne. Sådanne grammatikafsnit udfylder tydeligvis et behov, eftersom ordbøgerne vistnok sælger godt.

Så er der blot tilbage at overveje, om man i universalordbøgerne kunne strukturere grammatiktunge artikler bedre ved at forsyne ordbogen med en selvstændig grammatikdel, som der kunne gives referencer til i ordbogsartiklerne. Anvendt på de nævnte ordbøger ville en sådan integreret grammatik i ordbøgerne eksempelvis betyde:

- at Dudengrammatikkens valensmønstre også blev brugt i DUDEN-ordbøgerne.
- at valensmønstrene i POLITIKEN NUDANSK ORDBOG også blev brugt i ordbogsartiklerne.
- at reglerne for tyske præpositioners rektion fremgik både af ordbogsartiklerne for præpositioner og af et selvstændigt grammatikafsnit, og at der var krydshenvisninger fra ordbogsartikler til grammatikdelen i fx GADS DANSK/TYSKE ordbog og i fx GYLDENDALS nye DANSK/TYSKE ORDBOG.

Dermed ville man få integreret grammatiske informationer i ordbogsdelen og få præsenteret flere grammatiske informationer under opslagsordene.

Noter

- 1 Præpositioner slås relativt tit op i ordbøger, jvf. den store brugerundersøgelse, som H. E. Wiegand har gennemført og publiceret resultater fra i afhandlingen 'Fragen zur Grammatik in Wörterbuchbenutzungsprotokollen. Ein Beitrag zur empirischen Er-

- forschung der Benutzer einsprachiger Wörterbücher'. In: H. Bergenholtz, J. Mugdan (Hrsg.): *Grammatik und Lexikographie. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch...* Lexicographica. Series Maior. 1985. S. 20 - 98 (se s. 88 ff.). I dette i øvrigt udmærkede samleværk har Burkhard Schaefer i 'Zur Beschreibung der Präpositionen im einsprachigen deutschen Wörterbuch' (s. 278 - 307) beskrevet, hvorledes præpositioner defineres og praktisk behandles i nyere tyske ordbøger.
- 2 Se litteraturlisten for nyere danske og tyske ordbøger med selvstændige grammatik-afsnit.
 - 3 Udtrykket *universalordbog* vælger jeg her at forstå som en modsætning til *specialordbøger*, der gør én informationstype til hovedsagen, fx er retskrivnings-, udtale-, synonym-, valens-, frekvensordbøger eksempler på specialordbøger. På tysk taler vi om *Gesamtwörterbücher* og *Spezialwörterbücher*, jvf. Hans-Peder Kromann, 'Wörterbücher und ihre Benutzer. Wörterbücher mit Deutsch als Objektsprache'. In: V. Ågel/ R. Hessky (Hgg.) *Offene Fragen - offene Antworten in der Sprachgermanistik*. Tübingen 1992, 151 - 164.
 - 4 Indeks **V 800** er i ordbogens valensafsnit fastsat til at stå for **Subj. + Vb + Präpobj.** med typeeksemplet *seine Aussage beruht auf Wahrheit*. Læg mærke til, at her kan man ikke på eksemplet se, om *auf* styrer akkusativ eller dativ.
 - 5 Wahrig har hentet inspiration i engelsk leksikografi, der har en længere tradition for at bringe *verb patterns* i deres ordbøger for udlændinge.
 - 6 Præpositioner i tosprogsordbøger lader jeg ligge i denne sammenhæng, fordi kravet om ækvivalens medfører et par ekstra spørgsmål, som der ikke er plads til at diskutere her. I den lille artikel 'Zur Syntax im Übersetzungswörterbuch' In: *Kopenhagener Beiträge zur germanistischen Linguistik*. Sonderband 3, 1987, s. 142 - 150, samt i artikel 287 'Grammatical Constructions in the Bilingual Dictionary' (sammen med Th. Riiber, P. Rosbach) i *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie...* Dritter Teilband 1991, 2770 - 2775, har jeg fremlagt synspunkter på grammatiske informationer i tosprogsordbøger. Eksempler med præpositioner i tosprogsordbøger findes i Gitte Baunebjerg Hansen *Artikelstruktur im zweisprachigen Wörterbuch. Überlegungen zur Darbietung von Übersetzungsäquivalenten im Wörterbuchartikel*. Tübingen 1990. (Lexicographica. Series Maior 35), side 166ff og 184ff.
 - 7 Det vil her være for pladskrævende både at gengive og gennemgå hele ordbogsartiklen. Jeg har udvalgt et enkelt eksempel fra hver af de fremanalyserede betydninger, men har ikke gengivet selve betydningsangivelsen. Interesserede henvises til at studere de originale ordbogsartikler for præpositioner.
 - 8 I WAHRIG DEUTSCHES WÖRTERBUCH hovedinddeles *über*-artiklen ganske vist ud fra kasusrektion <Akkusativ> og <Dativ>, men der formuleres alligevel ikke nogen regel for, hvornår præpositionen styrer akkusativ og hvornår dativ.
 - 9 Begreberne intralokal og translokal er lånt fra P. Jørgensen *Tysk Grammatik I - III*. København 1992. (II, 103 - 174).

- 10 For en udførlig introduktion til 'hoved' og 'hale' i dansk leksikografi henvises til Gitte Baunebjerg Hansen (jvf. ovenfor, note 6), se især side 102 ff.
- 11 Der kan selvfølgelig - som vi har set det i WAHRIG-dtv - være en lille selvstændig syntaksdel med valensmønstre, som der så henvises i ordbogsartiklerne.

Ordbøger

Danske og tyske ordbøger med selvstændige grammatikafsnit

- 1993: *Dansk-Tysk Ordbog. Undervisning.* Af Grethe Hjorth. Sproglig konsulent Ingeborg Zint. Hovedredaktør Jens Axelsen. 1. udgave. Gyldendal, København. 800 s. Ingeborg Zint: *Grammatik.* S. 758 - 800.
- 1990 *Politikens Nudansk Ordbog.* Redaktion: Christian Becker-Christensen, Peter Widell. 14. udgave. København 1990 (¹1953) 1160 s. (Lars Heltoft:) *Grammatisk oversigt.* S. 47 - 68.
- 1989 *Deutsches Universalwörterbuch. 2., völlig neu bearb. u. stark erw. Aufl.* Hrsg. u. bearb. v. Wiss. Rat u. d. Mitarb. der Dudenred. unter Leitung von Günther Drosdowski. Dudenverlag. Mannheim, Leipzig, Wien, Zürich. 1989. 1816 s.
Kurze Grammatik der deutschen Sprache. S. 15 - 48.
- 1987 *Tysk-dansk, dansk-tysk.* Redaktion: Henning Poulsen. Konsulenter Hans-Peder Kromann, Monika Jensen. 1. udgave. Gad, København 1987. 823 s. Hans-Peder Kromann: *Tysk grammatik.* S. 789 - 823.
- 1978 *dtv-Wörterbuch der deutschen Sprache.* Hrsg. v. Gerhard Wahrig in Zus. m. zahlr. Wiss. u. a. Fachl. Deutscher Taschenbuchverlag, München 1978. 943 s.
Tabellen zur Formenbildung und Syntax. S. 14 - 33.
- 1966 Gerhard Wahrig: *Das große deutsche Wörterbuch.* Hrsg. in Zus. m. zahlr. Wiss. u. a. Fachl. Mit einem "Lexikon der deutschen Sprachlehre". Bertelsmann, Gütersloh 1966.
- 1975 *Lexikon der deutschen Sprachlehre:* Spalte 46 - 250. (Af 4321 spalter)
- 1937 *Dansk-Tysk Ordbog.* Udarbejdet af J.J. Siegler. 431 s. Berlingske Forlag, København.
J.J. Siegler: *Grammatik.* 417 - 429.

Strejflys over islandsk leksikografis historie.

Guðrún Kvaran

I mit foredrag vil jeg forsøge i grove træk at give en oversigt over islandsk leksikografis historie. Jeg vil for det meste holde mig til større ordbøger som optager såvel ældre som yngre islandsk ordforråd tilhørende forskellige genrer. Glossarier til enkelte værker omtales ikke, men derimod enkelte specialordbøger som behandler afgrænsede dele af ordforrådet. I enkelte tilfælde nævnes ordbøger over fremmedsprog, når deres islandske oversættelser udgør værdifulde bidrag til ordforrådets historie.

Man kan sige at islandsk leksikografi begynder med den nordlandske præst Magnús Ólafssons islandske glossarium. Det blev længe antaget at Magnús udførte sit arbejde på professor Ole Worms initiativ, men i sin brevveksling med Magnús omtaler Worm først 1635 det ønskelige i at der blev udarbejdet et glossarium over gamle ord og udtryk i islandsk poesi. Allerede året efter døde pastor Magnús og efterlod da manuskriptet til et ganske vist ufærdigt glossarium, som Worm fik tilsendt og udgav i udvidet form i København i 1650 med titelen *Specimen Lexici Runici*. Som følge af Worms specielle interesse for runer lod han alle opslagsord trykke med runer. Anthony Faulkes, som har undersøgt de kilder som ordbogen bygger på, gjorde opmærksom på at Magnús kun i ringe grad har excerperet poesi, og han formoder at Magnús allerede inden Worms henvendelse havde udarbejdet grundstammen til en ordbog. Det ordbogsmanuskript af Magnús ordbog, som Worm fik i hænde, er gået tabt, men af en sammenligning af en samtidig afskrift og udgaven fremgår at der er tale om væsentlige tilføjelser, til dels fra Worm og til dels fra den islandske student Guðmundur Andrésón som selv skrev en islandsk ordbog.

Guðmundur Andrésóns ordbog, *Lexicon Islandicum*, blev først udgivet i København 1683, næsten tredive år efter hans død. Professor Peder Hansen Resen besørgede udgaven, men som følge af hans begrænsede beherskelse af det islandske sprog er udgaven skæmmet af et utal fejllæsninger og trykfejl.

Som i *Specimen Lexici Runici* er ordforklaringerne i *Lexicon Islandicum* på latin. Man måtte dog længe lade sig nøje med disse to ordbøger samt Verelius' *Index lingvæ veteris Scytho-Scandicæ sive Goticæ* som blev udgivet posthumt af Rudbeck i Uppsala 1691, idet der gik over et hundrede år inden der udkom en ny islandsk ordbog, som jeg vil omtale senere. Savnet heraf fremgår f.eks. klart af Rasmus Rasks skrifter.

I *Vejledning til det islandske eller gamle nordiske Sprog* 1811 nævner Rask de to københavnske ordbøger og siger at de er „næsten aldeles ubrugbare“ og „Det lønner ikke Umagen at beskrive dem nøjere; men mærkeligt er det, at en Svensker har overgaaet

tvende Islændere i at samle en islandsk Ordbog.“ Set med nutidens øjne er Rasks vurdering af de to første ordbøger lovlig streng. Deres værdi ligger blandt andet deri, at *Specimen Lexici Runici* - som påpeget af Faulkes - giver oplysninger om islandsk talesprog i det 17. århundrede, og det samme gælder *Lexicon Islandicum*, at man der kan finde ord fra talesproget som ikke findes i senere ordbøger. For tiden forberedes en ny udgave på Orðabók Háskólans af *Lexicon Islandicum* med tilbørlig hensyntagen til et bevaret manuskript af værket.

Selv om der ikke udkom en ny islandsk ordbog før inde i det 19. århundrede, så blev der i mellemtiden arbejdet med islandske ordbogsprojekter i forbindelse med en stor interesse for den gamle islandske litteratur, både i Sverige og Danmark.

I Sverige arbejdede unge islændinge næsten kontinuerligt med oldskriftgranskning fra omkring 1660 til henimod år 1700, og til trods for deres berygtede drikfældighed fik de udrettet en del. En af disse islændinge var Guðmundur Ólafsson, som blandt andet blev sat til at udarbejde en islandsk-latinsk ordbog. Ved hans død i 1695 var arbejdet ikke færdigt, men det er bevaret som manuskript i 16 bind i Kungliga biblioteket i Stockholm. Senere leksikografer har næsten ikke anvendt dette manuskript, som længe var glemt. Efter min mening fortjener dette værk en systematisk udnyttelse. Et fotografisk eksemplar af denne ordbog har på Orðabók Háskólans tit været til nytte, f.eks. som den ældste kilde til senere kendt ord.

Betydelig mere bekendt er et andet ordbogsmanuskript, som er bevaret i ni kæmpestore bind i Den arnamagnæanske samling. Forfatteren var Jón Ólafsson fra Grunnavík, der som student blev Árni Magnússons assistent få år før hans død og levede som arnamagnæansk stipendiat i København det meste af sit voksne liv. Jón Ólafsson påbegyndte sit værk i 1734 og blev ved med at udvide det til sin død i 1779.

I denne ordbog er der også oversættelser og forklaringer på latin, og i visse tilfælde også på dansk eller islandsk. Ordforrådet omfatter såvel oldsproget som senere islandsk sprog, og som et kulturhistorisk leksikon er det en meget betydningsfuld kilde. En hovedgrund til at håndskriftet ikke er blevet udgivet er formodentlig den, at på grund af de mangfoldige tilføjelser - mellem linjerne, i marginen og på løse sedler - er håndskriftet svært at have med at gøre. Professor Jón Helgason har udtalt at hvis ordbogen var blevet ordentlig redigeret ville man ufortøvet betegne den som en af de største islandske lærdomspræstationer i det 18. århundrede. Foruden ord hentet fra skrevne og trykte kilder indeholder Jón Ólafssons ordbogsmanuskript et meget stort antal ord som han har hentet fra samtidens talesprog. Af den grund har man her i mange tilfælde de ældste belæg på ord som ikke er fundet ved excerpering af ældre bøger, således at værket giver værdifulde oplysninger om islandsk ordforråds historie. Det anvendes derfor i stigende grad af leksikografer og er for længe siden excerperet for Orðabók Háskólans af ordbogens mangeårige chef Jakob Benediktsson.

En anden vigtig kilde til islandsk ordforråd i det 18. århundrede er en latinsk-islandsk ordbog, *Nucleus Latinitatis*, som biskop Jón Árnason oversatte og udgav i 1738. Forbilledet er en dansk bog af samme navn ved Hans Gram, udarbejdet i 1709-1710, men den ældste bevarede udgave er fra 1722. En sammenligning med det danske værk viser at bi-

skop Jón har overtaget den latinske ramme, d.v.s. opslagsord og citater, og forsynet dem med islandske oversættelser. Bogen var naturligvis beregnet til brug i de islandske katedralskoler og forfatteren har bestræbt sig for ikke alene at klargøre opslagsordenes betydning, men samtidig at give flere alternativer med hensyn til oversættelse, således at ordbogen faktisk i mange tilfælde også fungerer som en islandsk synonymordbog. Man kan således antage at bogen kan have påvirket peblingernes sprogbrug.

I Grams *Nucleus* efterfølges den latinsk-danske del af bogen af en alfabetisk liste over de danske ord som anføres som oversættelser med henvisning til side og linje. Det er ikke usandsynligt at biskop Jón Árnason oprindeligt har tænkt sig at have en tilsvarende islandsk ordliste i sin bog, men at han har opgivet tanken og i stedet bestemt sig til at udgive en islandsk-latinsk ordbog. Et sådant værk foreligger i manuskript i Landsbókasafn Íslands (Landesbiblioteket), for det meste med biskoppens egen hånd, og det antages at være skrevet omkring 1740. Det er et voluminøst værk, næsten 1200 sider. Ordforrådet består for det meste af almindelige ord, og ordvalget tyder på at værket har været tænkt som et hjælpemiddel ved latinsk stilskrivning (Biskop Jón var en dygtig pædagog og er kendt for sin interesse for reformer af det islandske skolevæsen.)

En af de elever som fik sin uddannelse på Skálholt i biskop Jón Árnasons tid, lige efter at hans *Nucleus Latinitatis* var udkommet, var den senere præst Björn Halldórsson, som selv til sin tid kom til at gøre en stor indsats på det leksikografiske område. Han har sandsynligvis påbegyndt sin islandsk-latinske ordbog omkring 1770, og 1786 rejste han til København med en renskrift af hele værket som omfattede næsten 30.000 ord. Her tilbød han Den arnamagnæanske kommission ordbogen, som købte den for 100 rigsdaler. Imidlertid trak udgivelsen i langdrag af forskellige grunde, blandt andet den at ordforrådet ansås for at være for begrænset, især med hensyn til den ældre poesis vokabularium. Først ca 25 år efter at Björn Halldórssons manuskript var kommet kommissionen i hænde kom der gang i sagerne, da to bemedlede nordmænd påtog sig at skaffe penge til udgivelsen. Samtidig blev Rasmus Rask antaget som redaktør og til at supplere pastor Björns latinske oversættelser med tilsvarende danske. Bogen udkom i København 1814 med titelen *Lexicon Islandico-Latino-Danicum*, godt 1000 sider.

Det er klart at Björn Halldórsson må have anvendt en latinsk ordbog ved udarbejdelsen af sin ordbog. Det mest nærliggende er naturligvis at antage at dette har været *Nucleus*, og en flygtig sammenligning af de to værker tyder også på det. Fra *Nucleus* er også sikkert en del af det islandske ordforråd hentet. Noget af det stammer fra ældre bøger, men største delen har uden tvivl været levende talesprog. Björn Halldórssons ordbog er således - ved siden af *Nucleus Latinitatis* - en vigtig kilde til islandsk ordforråd i det 18. århundrede, og i dem har Orðabók Háskólans i mange tilfælde de ældste belæg på ord og ordforbindelser som man kender fra det senere sprog. Bogen blev genudgivet af Orðabók Háskólans, og her arbejder man nu med en genudgivelse af *Nucleus Latinitatis* som skal suppleres med en islandsk ordliste.

Grunden til at Orðabók Háskólans er gået i gang med genudgivelse af gamle ordbøger er den, at de er en vigtig kilde til tidligere århundreders ordforråd, og deres anvendelse er betydelig lettere når de gøres tilgængelige i maskinlæsbar form; dette sidste gælder

naturligvis primært ordbøger over fremmedsprog med islandske oversættelser, hvor det islandske materiale kan alfabetiseres maskinelt. Desuden har det vist sig at der er et marked for disse genudgivelser, idet kun få interesserede har været i besiddelse af originaludgaverne.

Björn Halldórssons ordbog har haft betydelig indflydelse på senere islandsk leksikografi. Først kan vi nævne en dansk-islandsk ordbog ved den arnamagnæanske stipendiat, senere domkirkepræst i Reykjavík, Gunnlaugur Oddsson, der udkom 1819 og blev genudgivet med islandsk ordliste af Orðabók Háskólans 1991. Som det fremgår af bogens titel indeholder den sjældne, fremmede og sværtforståelige ord der forekommer i danske bøger. Ordbogen er altså primært en dansk fremmedordbog med islandske oversættelser, selvom den også indeholder en del sjældne danske ord, dels gamle og dels nydannelse. Et forbillede er sikkert danske fremmedordbøger fra århundredets begyndelse, hvis formål først og fremmest var puristisk. Dette har udgiveren af 1991-udgaven, min kollega Jón Hilmar Jónsson, påpeget, og samtidig har han påvist at Gunnlaugur Oddsson har haft intimt kendskab til Björn Halldórssons ordbog og herfra hentet adskillige yderst sjældne ord som han anvender som oversættelser i sin bog.

En anden islandsk leksikograf som hentede materiale i Björn Halldórsson ordbog var professor Konráð Gíslason i København. Han udsendte en dansk-islandsk ordbog i 1851 efter at have arbejdet med den ca. ti år med afbrydelser. Konráðs forbillede var *Dansk Ordbog* af Christian Molbech, som udkom i København 1828-33. Konráð Gíslason var en af de islændinge som på denne tid gik stærkt ind for purisme, og med sin bog ville han bidrage til for sine landsmænd at lette brugen af islandske ord i stedet for dansker. Imidlertid havde han ikke meget tilovers for neologismer, således at når han ikke kunne finde en god gammel islandsk modsvarighed til et dansk ord foretrak han at beskrive betydningen fremfor at anføre et ord han ikke var tilfreds med. Det ser ikke ud til at denne ordbog har vakt nogen større opsigt da den udkom, men den har uden tvivl været til nytte for mange blandt andet i betragtning af at der gik næsten et halvt århundrede inden den næste dansk-islandske ordbog udkom, og også i dette århundrede vides at omhyggelige stilister som Halldór Laxness har gjort flittig brug af Konráðs ordbog.

Inden Konráð Gíslason tog fat på sin dansk-islandske ordbog havde han i nogle år arbejdet for en engelsk rigmand, Richard Cleasby, på en oldislandsk ordbog med engelske oversættelser. Da Cleasby døde synes den islandske del at have været nogenlunde færdig, men uenighed mellem Cleasbys arvinger og Konráð resulterede i at materialet blev sendt over til England, hvor det henlå en halv snes år, indtil en anden islænding blev ansat til at overtage arbejdet. Det var Guðbrandur Vigfússon i Oxford som gjorde udgaven færdig, og i fortalen bagatelliseres Konráð Gíslasons andel i bogen. Den udkom i hæfter 1869-74 under titelen *An Icelandic-English Dictionary*, men går som regel under navnet *Cleasby-Vigfússon*.

Selv om bogen først og fremmest skulle omfatte oldislandsk ordforråd, så indeholder den derudover et ret tilfældigt udvalg af ord og ordforbindelser fra senere islandsk, som Guðbrandur Vigfússon har tilføjet.

Kort inden Cleasby-Vigfússons ordbog udkom havde et par andre ordbøger set da-

gens lys. Den første af disse var rektor Sveinbjörn Egilssons ordbog over den norrøne poesi, *Lexicon Poeticum*, som blev trykt i København 1854-60. Oprindeligt samlede Sveinbjörn også materiale fra prosasproget, men på Carl Christian Rafns opfordring indskrænkede han sig til det poetiske sprog for ikke at kolliderede med Cleasbys planlagte ordbog. Sveinbjörn skrev sine ordforklaringer på latin, men i senere udgaver af værket, besørget af professor Finnur Jónsson, blev latinen afløst af dansk.

Andre oldnordiske eller norrøne ordbøger der kom før Cleasby-Vigfússon var *Oldnordisk ordbog*, København 1863, ved Eiríkur Jónsson, som i sin tid havde arbejdet sammen med Konráð Gíslason på Cleasbys ordbogsprojekt, og første udgaven af *Ordbog over det gamle norske sprog*, Kristiania 1867, ved den norske præst Johan Fritzner. Begge disse ordbøger var praktiske håndbøger for dem der ville læse islandske og andre norrøne middelaldertekster.

Den mest omfattende ordbog på dette område er anden udgave af Fritzners ordbog som udkom i tre bind 1883-1896 med et væld af citater, hentet blandt andet fra det store antal middelalderlige tekster som blev udgivet - mange for første gang - i sidste halvdel af det 19. århundrede.

Fra islandsk side var det vigtigste bidrag til leksikografien i det 19. århundredes sidste fjerdedel rektor Jón Þorkelssons *Supplement til islandske Ordbøger* der udkom i fire bind i Reykjavík 1876-99. Bind 1 og 2 er først og fremmest supplement til første udgaven af Fritzners ordbog samt Cleasby-Vigfússon, og indeholder desuden en del materiale fra det 16. og 17. århundrede. Bind 3 er tænkt som et supplement til Björn Halldórssons ordbog og bygger på excerpering af en del islandske bøger der udkom mod slutningen af det 18. og i det 19. århundrede. Dette bind indeholder mange nye ord som var blevet til som følge af erhvervsmæssig fremskridt og øget bogproduktion. I bind 4 er materialet igen for det meste hentet fra middelalderlige kilder, men her lægges hovedvægten på formværket.

Skønt en del større og mindre ordsamlinger ikke er nævnt, skulle det af denne oversigt fremgå, at der i løbet af forrige århundrede blev udført stort arbejde på det leksikografiske område. De fleste af de ordbøger som her er nævnt blev udarbejdet af idealister i deres fritid og til dels på grundlag af begrænsede hjælpemidler. I helhed betragtet viser de kontinuiteten inden for islandsk leksikografi.

På denne grund byggede bibliotekar Sigfús Blöndal da han i begyndelsen af dette århundrede i København skrev sin islandske ordbog med danske oversættelser som blev trykt i Reykjavík i årene 1920-1924. Han benyttede sig af ordforrådet i Björn Halldórssons ordbog og Jón Þorkelssons *Supplementum*, men desuden excerperede han og hans medarbejdere et stort antal trykte bøger og indsamlede ord og udtryk fra talesproget. Blöndals ordbog var langt den udførligste kilde til islandsk sprog i de senere århundreder, og den har stadig ikke fået en afløser. Den er blevet optrykt to gange og fik i 1963 et supplement med ord fra det 20. århundrede.

Det samme år udgav lektor Árni Böðvarsson for Menningarsjóður (det nu nedlagte statsforlag) den første islandske ordbog med islandske forklaringer - bortset fra et tilløb ved redaktør Jón Ólafsson i 1912-15, som kun nåede frem til *br* i alfabetet. I revideret og

udvidet form er Árni Böðvarsson's ordbog stadig den eneste ordbog af sin slags på markedet, og yderligere revision af den er nu under forberedelse.

Ved den tid da Blöndals ordbog udkom blev der fremsat tanker om at udarbejde en historisk ordbog som omfattede hele den islandske sprogperiode. Sagen blev diskuteret, og arbejdet kom så småt i gang, men det blev snart opgivet igen på grund af svigtende bevillinger. Resultatet blev dog at man i fyrreerne besluttede at gå i gang med en ordbog som dækkede perioden fra 1540 op til nutiden.

Excerpering til en historisk ordbog over islandsk sprog efter 1540 er nu foregået på Orðabók Háskólans i næsten et halvt århundrede. Her er der ikke tid til at beskrive dette ordbogsarbejde, som i de senere år er blevet meget forandret, dels om følge af at ordbogen er blevet defineret som et leksikografisk institut og ikke mindst fordi datamaskiner i høj grad har ændret arbejdsformen. Her skal det lige nævnes at der nu foreligger det første prøvehæfte, som er baseret på den redaktion af verber som i de sidste år er foregået på ordbogen.

Jeg har her ikke kunnet omtale en del gode ordbøger der er kommet inden for det sidste årti, men jeg vil dog til slut nævne en stor engelsk-islandsk ordbog der kom i 1984, og en dansk-islandsk ordbog, baseret på *Nudansk ordbog*, 1992 og ordbogsredaktør Ásgeir Blöndal Magnússon's etymologiske ordbog der blev udgivet 1989.

I det foregående har jeg forsøgt at give en oversigt over islandske ordbøger op til vor tid, d.v.s. såvel ordbøger over det islandske sprog som nogle fremmedordbøger med islandske oversættelser. For en leksikograf, som arbejder på en historisk ordbog, er det meget vigtigt at kende kontinuiteten i leksikografien og forstå hvorledes det ene ordbogsværk bygger på det andet fra de ældste ordbøger og op til nutiden.

Litteratur:

- Árni Böðvarsson. 1963. *Íslensk orðabók handa skólum og almenningi*. Reykjavík. (Anden udgave forøget, Reykjavík 1983).
- Ásgeir Blöndal Magnússon. 1989. *Íslensk orðsifjabók*. Reykjavík.
- Björn Halldórsson. 1814. *Lexicon Islandico-Latino-Danicum Björmonis Haldorsonii*. Vol. I-II. Havniæ.
- Björn Halldórsson. 1992. *Orðabók. Íslensk-latenesk-dönsk*. Eftir handriti í Stofnum Árna Magnússonar í Kaupmannahöfn. Ny útgáfa. Jón Aðalsteinn Jónsson sá um útgáfunna. Reykjavík.
- Eiríkur Jónsson. 1863. *Oldnordisk Ordbog ved det kongelige nordiske Oldskrift-Selskab*. Kjöbenhavn.
- Faulkes, Anthony. 1964. The Sources of Specimen Lexici Runic. *Lingua Islandica - Íslensk tunga*. 5:30-138. Reykjavík.
- Fritzner, Johan. 1867. *Ordbog over det gamle norske Sprog*. Kristiania.
- Guðbrandur Vigfússon. 1874-76. *An Icelandic-English Dictionary based on the MS. Collections of the late Richard Cleasby enlarged and completed by Gudbrand Vigfusson*. With an Introduction and Life of Richard Cleasby and George Webbe Dasent. Ox-

- ford. (2nd Edition. With a Supplement by Sir William Craigie. Oxford 1957).
- Guðmundur Andr sson. 1683. *Lexicon Islandicum Sive Gothica Runa vel Lingva Septentrionalis Dictionarium ...* Hafniae.
- Gunnlaugur Oddsson. 1819. *Or ab k, sem inniheldur flest f gi t, framandi og vandskilinn or , er ver a fyrir i danskum b kum.* Kaupmannah fn.
- Gunnlaugur Oddsson. 1991. *Or ab k, sem inniheldur flest f g t, framandi og vandskilinn or , er ver a fyrir i d nskum b kum.* Ny  tg fa med  slenskri or askr . J n Hilmar J nsson s  um  tg funa  samt Þord si  lfarsd ttur. Reykjav k.
- Hrefna Arnalds, Ingibj rg Johannesen (red.) 1992. *D nsk- slensk or ab k.* Reykjav k.
- J n  rnason. 1734. *Nucleus latinitatis, Qv  pler qve Romani sermonis Voces, ex classicis Auctoribus aure  argente qve  tatis, ordine Etymologico adduct , & Interpretatione vernacula exposit  comprehenduntur.* Hafniae.
- J n Helgason. 1926. *J n  lafsson fr  Grunnav k.* Safn Fr  afjelagsins um  sland og  slendinga. V. Kaupmannah fn.
- J n Þorkelsson. 1876-99. *Supplement til islandske ordb ger. I-IV. Samling.* Reykjav k og Kaupmannah fn.
- Konr   G slason. 1851. *D nsk or ab k med  slenskum þy ingum.* Kaupmannah fn.
- Magn s  lafsson. 1650. *Specimen Lexici Runici, Obscuriorum qvarundam vocum, qv  in priscis occurrunt Historiis & Po tis Danicis, enodationem exhibens ...* Hafniae.
- Rask, Rasmus Kristian. 1811. *Vejledning til det Islandske eller gamle nordiske sprog.* K benhavn.
- Sigf s Bl ndal. 1920-24. * slensk-d nsk or ab k.* Reykjav k.
- Sveinbj rn Egilsson. 1860. *Lexicon Po ticum antiqu  lingua Septentrionalis.* Hafniae. (2. udgave ved Finnur J nsson, K benhavn 1913-1916).
- S ren S rensson. 1984. *Ensk- slensk or ab k me  alfr  ilegu  vafi.* Reykjav k.
- Verelius, Olaf. 1691. *Index lingv  veteris Scytho-Scandicae sive Gothicae ex vetusti  vi monumentis, maximam partem manuscriptis, collectus atqve opera Olai Rudbecki editus.* Upsal .

Skandinavisk ordbok – rapport från ett pågående projekt

Birgitta Lindgren

Syftet

Syftet med vårt projekt är att framställa en skandinavisk ordbok som tar upp sådana ord som kan vålla problem mellan danskar, norrmän och svenskar.

Initiativet har kommit från det svenska bokförlaget Norstedts, och boken kommer att ges ut av detta förlag i samverkan med det danska Gyldendal och det norska Kunnskapsforlaget tillsammans med Nordiska språksekretariatet.

Dansk redaktör är Allan Karker, norska redaktörer är Skirne Bruland och Ståle Løland och själv är jag svensk redaktör och svarar för den praktiska samordningen. Vi räknar med att vara klara nästa år.

Jag vill betona att det är ett pågående arbete som presenteras. De exempel som anförs skall demonstrera principer och riktlinjer, och skall inte uppfattas som utdrag ur en färdig ordbok.

Andra skandinaviska ordböcker

Den äldsta skandinaviska ordboken är "Svensk-dansk-norsk ordbog" från 1911 av Ida Falbe-Hansen. Det finns också en "Samnordisk ordbog" av Andreas Fynning från 1958 (2 uppl. 1971) med en mycket speciell uppläggning. Den gör verkligen skäl för sitt namn, för den förtecknar det gemensamma ordförrådet – kanske lite väl entusiastiskt i-bland (se nedan) – och utelämnar det särspråkliga, dvs. det rakt motsatta till den ordbok vi tänker framställa.

(ur Fynning "Samnordisk ordbog")

| <u>danska</u> | <u>norska</u> | <u>svenska</u> |
|-------------------|-------------------|-------------------|
| <i>adelskab</i> | <i>adelskap</i> | <i>adelskap</i> |
| <i>adelsmærke</i> | <i>adelsmerke</i> | <i>adelsmärke</i> |
| <i>adelsslægt</i> | <i>adelsslekt</i> | <i>adelssläkt</i> |
| <i>adelsstand</i> | <i>adelsstand</i> | <i>adelsstånd</i> |
| ... | ... | ... |
| <i>beat</i> | <i>beat</i> | <i>beat</i> |
| <i>Beatles</i> | <i>Beatles</i> | <i>Beatles</i> |

Sedan kan nämnas att det finns vissa smärre skandinaviska ordlistor, t.ex. den samordnade listan i "Att förstå varandra i Norden" samt specialiserade ordlistor, som t.ex. "Nordisk förvaltningsordbok".

Därutöver finns givetvis tvåspråkiga ordböcker och ordlistor, t.ex. de så kallade *basordlistorna* utgivna av Nordiska språksekretariatet och bokklubben Norden på 80-talet, som dock inte innehåller mer än några hundratal ord, samt de större ordböckerna, t.ex. "Norstedts dansk-svenska ordbok" (3 uppl. 1980) eller "Norstedts norsk-svenska ordbok", som kom ut förra året. Det skall sägas att vår skandinaviska ordbok på intet sätt gör anspråk på att konkurrera med sådana större ordböcker.

Användaren

Den tilltänkta användaren är alltså en person som har behov att kunna slå upp ord på något av de två grannspråken, inte för att åstadkomma en översättning på det egna språket utan för att få en förklaring som ger honom tillräckliga upplysningar för att förstå ordet i sitt sammanhang. De texttyper vi tänker oss är modern skönlitteratur, tidningstext och annan allmän sakprosa.

Urval och omfång

Vi tar inte med ord som är helt lika, t.ex. *arv*, *bibliotek*, *stål*, eller sådana som exemplifierar lätt igenkännliga systematiska skillnader, t.ex. rent grafiska skillnader som i *bær/bår*, *brød/bröd*, *ikke/icke*, skillnader i ändelser som i *leve/leva*, skillnader som i *pæls/päls*, *gab/gap*, *ud/ut*, *sag/sak*.

Vid urvalet håller vi oss inom den del av ordförrådet i varje språk som kan sägas utgöra det allmänna, vanliga ordförrådet, dvs. sådana ord som hamnar bland de mest frekventa i frekvensundersökningar eller ord som uppfattas som vanliga och välkända, även om de inte råkat komma med i publicerade frekvensundersökningar. Vi undviker alltså ovanliga, ålderdomliga, alltför snävt fackspråkliga ord. Vi är också relativt restriktiva gentemot slangord.

Vi har satt en övre gräns vid 3 000 uppslagsord per språk, alltså sammanlagt ger 9 000 uppslagsord totalt, vilket vi bedömer skulle täcka den avsedda delen av ordförrådet.

Typ av ordbok

Till sin typ blir det en passiv ordbok. Utgångspunkten vid val av ord och utformning av artiklarna har hela tiden varit att användaren bara slår upp de ord i grannspråken som han inte förstår, när han stöter på dem i någon text, och då bara behöver en kortfattad förklaring, helst i form av en motsvarighet.

Detta innebär att vi inte ger upplysningar om uppslagsordens uttal, böjning eller konstruktionsmöjligheter, och inte heller anger om ett norskt uppslagsord tillhör bokmål eller nynorska.

Tilläggsinformation

Förutom själva ordboksdelen är det tänkt att ordboken också skall innehålla tillägg i form av översikter eller tabeller som visar på systematiska skillnader i uttal, stavning, ordböjning m.m. mellan de skandinaviska språken.

Disposition

Ordboken redigeras så att den delas upp på tre separata trespråkiga listor, alltså en lista med danska uppslagsord avsedd för norsk- och svensktalande, en lista med norska uppslagsord avsedd för dansk- och svensktalande, en lista med svenska uppslagsord avsedd för dansk- och norsktalande.

Andra alternativ hade varit att göra sex separata tvåspråkiga listor inom en enda volym eller att göra en enda samordnad lista. Det finns fördelar och nackdelar med alla lösningarna.

Att ha separata tvåspråkiga listor skulle ha fördelen att man vid urvalet av ord och utformning av förklaringar bara hade att ta hänsyn till ett målspråk, nackdelen är att man förlorar den pedagogiska finessen med trespråkiga listor, alltså att användaren utöver uppgiften om vad ett visst ord på källspråket motsvarar på hans språk också får veta vad det heter på det andra gransspråket. Troligen skulle en lösning med sex separata listor i en volym bli det mest utrymmeskrävande.

Att ha en enda lista vore naturligtvis en tilltalande lösning, just för att det är en enda lista. Det skulle troligen också vara det mest platssparande. För den enskilda användaren är det dock mindre praktiskt med en enda lista. Bortsett från att svenskar ibland kan ha svårt att skilja *skrivnen* danska från *skrivnen* norska, torde ändå normalfallet vara att man vet vilket språk en viss text är skriven på, och då är det bra att kunna gå till en lista som behandlar just det språket.

Finska och isländska motsvarigheter

De valda urvalsprincipen att bara ta med ord som är svåra i det interskandinaviska umgänget ställer stora krav på andra nordbor som vill använda ordboken. Från början ingick i våra planer att som en hjälp lägga till förklaringar på finska och isländska, men tills vidare har vi skrinlagt den delen.

Utformning av artikel

En artikel är alltså utformad så, att efter uppslagsordet kommer uppgifter om motsvarigheter på gransspråken. Under inskrivningsskedet ges dessa på ny rad för varje språk. Den slutliga uppställningen och grafiska utformningen är ännu inte avgjord. Som tidigare sagts ges inga upplysningar om uttal eller böjning, däremot kan böjningsformer som bedöms vara svårbegripliga på något av målspråken tas med som egna uppslagsord. Så t.ex. i den norska listan tar vi med presensformen *kjem*, vilket motsvaras av danska och svenska *kommer*.

Birgitta Lindgren

Förklaringar

Som förklaring väljer vi det vanliga, normala uttrycket på varje målsspråk, inte något uttryck som är språkhistoriskt identiskt men mindre vanligt. Som svensk motsvarighet till exempelvis danska *buntmager* ger vi *körsnär* och avstår från att tillfoga *buntnakare*, eftersom det ordet torde vara okänt för de flesta svenskar. Till *danne* ger vi *bilda* som motsvarighet och avstår från *dana*, eftersom *dana* har en så speciell användning i svenskan.

För att disambiguera flertydiga ord i förklaringsdelen har vi satt in synonymer. T.ex. till den svenska förklaringen *barnsäng* under danska *barsel* lägger vi *nedkomst*. Vi är emellertid öppna för tanken att stryka dessa tillägg, eftersom användaren utifrån sammanhanget bör kunna avgöra vilken betydelse som avses.

Uppslagsord med viss stilvalör får ingen uppgift om stiltillhörighet, däremot kan vi ge direkt motsvarighet på samma stilnivå när sådan finns. Ett exempel:

hvid: koste det hvide ud af øjnene

no: koste flesk

sv: kosta skjortan

Om uttrycket i förklaringen är flertydigt sätter vi in en synonym för att disambiguera det:

strømer

no: purk

sv: snut, polis

Om direkta eller näraliggande motsvarigheter saknas, ger vi en kortfattad förklaring:

hf (højere forberedelseseksamen)

no: en slags eksamen på den videregående skolen

sv: en typ av gymnasieexamen

Som redan framgått tar vi upp flerordiga uttryck som egna uppslagsord. Ytterligare ett exempel:

fisk: gå i fisk

no: falle i fisk

sv: gå i stöpet

Lömska ord

"Lömska ord", ord som alltså ser identiska ut med eller liknar ord i det egna språket men betyder något annat, eller som har ytterligare en betydelse som avviker från en gemensam betydelse, markeras med ett varningstecken, i manus dubbelt utropstecken respekt-

tive nummertecken (#). Vi anger också vad det ord på målspråket som liknar uppslagsordet motsvarar på källspråket. Exempel:

afdrag !!
no: avdrag
sv: avbetalning
(sv. *avdrag* = da. *fradrag*)

and #
no: and
sv: 1. and 2. anka

Det kan diskuteras om vi inför utgivningen skall behålla dessa uppgifter, eftersom sådant väl mera hör hemma i en aktiv ordbok.

Problem

Givetvis erbjuder också framställningen av en sådan här ordbok problem, ibland sådana som alla ordboksredaktörer kämpar med, ibland sådana som är speciella just för flerspråkiga ordböcker, i synnerhet för just den här.

När det gäller *urvalet* skall vi alltså bedöma om ordet tillhör det allmänna ordförrådet och om det kan anses vara svårbegripligt. Jag kan visa några exempel på danska ord som vi tänkt utelämna.

| | |
|--------------------|--|
| <i>drager</i> | (sv. <i>bärare</i>) inaktuelle fenomen |
| <i>fejekost</i> | onödig sammansättning, <i>feje</i> (sv. <i>sopa</i>) och <i>kost</i> (sv. <i>kvast</i> , <i>borste</i>) med som egna uppslagsord |
| <i>hakkebøf</i> | (ungefär sv. <i>pannbiff</i>), <i>hakkebøf</i> måste bedömas som genomskinligt ord för svenskar |
| <i>skovjordbær</i> | onödig sammansättning, <i>skov</i> med som eget uppslagsord, under <i>jordbær</i> anges att det motsvarar sv. <i>jordgubbe</i> och <i>smultron</i> |

Det inbördes förhållandet mellan de skandinaviska språken kan lite generaliserat beskrivas så, att danska och norska – i synnerhet bokmål – grupperar sig på en sida, och svenska – ibland i sällskap med nynorska – på den andra sidan. Det kan möjligtvis ge ett enformigt intryck.

Förhållandet ser alltså i schematiserad form ofta ut så här i den danska listan:

Birgitta Lindgren

A
no: A1
sv: B

A
no: A1, B
sv: C

A
no: A1, B
sv: B1

eller med exempel:

ædru
no: edru
sv: nykter

edderkop
no: edderkopp, kongro
sv: spindel

bevidst
no: bevisst, medveten
sv: medveten

Även här kan man naturligtvis fundera över hur detta förhållande skall framställas rent grafiskt i den tryckta ordboken.

Att ha *mer än ett målspråk* i en lista innebär vissa problem. Eventuella redaktionella kommentarer till ett uppslagsord bör då helst placeras i förklaringsdelen skrivna på målspråken, hellre än vid uppslagsordet skrivna på källspråket.

køer
no: (pl.) kyr
sv: (pl.) kor

Vidare angavs ju det vara ett pedagogiskt syfte med den valda uppställningen, att användaren samtidigt kan få se vad ett ord på det ena grannspråket heter på det andra grannspråket. Eftersom vi emellertid valt att ge så kortfattade upplysningar som möjligt, kan det kanske ibland te sig förbryllande:

dompap
no: dompap
sv: 1. domherre 2. dumbom

En svensk läsare kan här undra om redaktörerna glömt bort att sätta in en andra betydelse i den norska förklaringen. Men förhållandet är det att även det norska *dompap* motsvarar helt det danska ordet, dvs. det kan användas i överförd bemärkelse, om en dum person. Följande exempel från den norska listan kanske i förstone ser ännu mer förbryllande ur:

bløt #
da: 1. våd 2. blød
sv: 1. blöt 2. mjuk 3. tokig, dum

Även här har målspråken olika antal betydelser (av samma skäl som vid föregående exempel), men därtill kommer att "samma" ord som uppslagsordet förekommer i förkla-

ringarna i målspråken (*blød, blöt*) fast på olika platser. Men inte heller detta beror på slarv i redigeringen. Betydelse 1 i danska motsvaras av betydelse 1 i svenska, betydelse 2 i danska motsvaras av betydelse 2 och 3 i svenska.

Litteratur

- Att förstå varandra i Norden. Språkråd till nordbor i nordiskt samarbete.* Nordiska språksekretariatets skrifter 6. Oslo 1989
- Dansk-norsk ordliste.* Bokklubben Norden och Nordiska språksekretariatet. Oslo 1989
- Dansk-svensk ordlista.* Bokklubben Norden och Nordiska språksekretariatet. Oslo 1988
- Nordisk förvaltningsordbok.* Rev. uppl. Nordiska rådet. 1991
- Norsk-dansk ordliste.* Bokklubben Norden och Nordiska språksekretariatet. Oslo 1989
- Norsk-svensk ordlista.* Bokklubben Norden och Nordiska språksekretariatet. Oslo 1988
- Norstedts norsk-svenska ordbok.* Utarbetad av Birgitta Fowler och Kjell Ivar Vannebo i samarbete med Lena Moberg vid Svenska språknämnden. Stockholm: Norstedts 1992
- Norstedts dansk-svenska ordbok.* Utarbetad av Bertil Molde under medverkan av Niels Ferlov. 3 uppl. Stockholm: Norstedts 1980
- Sammordisk ordbog/ordbok.* Red. Andreas Fynning. 2 uppl. København: Nyt Nordisk Forlag 1971
- Svensk-dansk ordliste.* Bokklubben Norden och Nordiska språksekretariatet. Oslo 1989
- Svensk-norsk ordliste.* Bokklubben Norden och Nordiska språksekretariatet. Oslo 1989

Svensk ordbok 1948

Carl-Erik Lundbladh

År 1948 gav det som då hette Svenska dövstumläraresällskapet ut en illustrerad svensk ordbok, som kallades Svensk ordbok. Den distribuerades genom dövstumläraresällskapet till ett pris av fyra kronor för ”dövstumma” och ”dövstumlärare”. Övriga köpare fick betala 20 kronor. Ordboken användes (och används än) i dövskolorna och delades ut gratis till barnen i avgångsklassen. Däremot såldes den inte mycket via bokhandeln.¹

Svensk Ordbok -48 blev i allmänhet okänd i kretsar som inte har med döva att göra. Mig veterligt finns den inte medtagen i någon översikt av svenska ordböcker, i varje fall inte i Sigurd 1968 eller Wörterbücher 1990. Det är orätt, men framför allt är det synd att den inte har använts i vanliga skolor eller av allmänheten och sannolikt i för liten eller ingen utsträckning av dem som skrivit andra svenska ordböcker. För att sprida mer kännedom om Svensk ordbok från 1948 beskrivs här översiktligt dess storlek, uppställning och sätt att redovisa ordens form och betydelse. Först emellertid något om tillkomsten.

1927 tillsattes en kommitté bestående av dövläraerna G. Hartman, Lund, J. Johansson, Manilla, och O. Ahlner, Gävle. G. Hartman var kommitténs ordförande.² När han gick i pension 1942 efterträddes han av J. Johansson som rektor för Östervångskolan i Lund. Denne färdigställde ordboken.³

Av ordbokens förord framgår att dess redaktion anlitat tre lundensiska språkmän och lexikografer såsom granskare, nämligen förre chefen för SAOB, professor Ebbe Tuneld, förre redaktören vid samma ordbok, aktuarien N. R. Palmlöf samt professor K. G. Ljunggren (som också varit redaktör vid SAOB). Redan 1931 hade Ebbe Tuneld avgivit ett utlåtande om ordboken, där han bl. a. säger att den ”är ett alldeles förträffligt arbete. Det medtagna ordmaterialet är synnerligen rikhaltigt. Översättningarna (definitionerna) äro korrekta och avfattade i en så enkel form, att de lätt kunna förstås. Uttalsuppgifter och belysande språkprov ha i stor mängd medtagits. Den utgör ett svenskt handlexikon, som med fördel skulle kunna användas av andra än dem, för vilka det närmast är avsett.”² Detta är nog inte bara vänliga överord. Åtminstone påpekandet att ordboken med fördel skulle kunna användas av andra än döva måste ha varit en sanning då och gäller i viss mån fortfarande, trots att Sverige nu har en så bra och genomarbetad definitionsordbok som namnen Svensk ordbok från 1986.

Omfång och utformning

Svensk ordbok -48 är på 976 tvåspaltiga sidor. Antalet huvuduppslagsord, som har halv-

fet stil och inleder rad med indrag, är enligt översiktlig beräkning ca 13.400. Därtill kommer minst lika många underordnade uppslagsord, mest sammansättningar och avledning, i spärrad rak stil. Dessa står i stycke som inleds av ett semantiskt och morfologiskt besläktat huvuduppslagsord. Det totala antalet uppslagsord torde alltså vara nästan 27.000. Det kan verka mycket i jämförelse med Svensk ordbok -86, som har 60.000, men den senare är restriktiv vad gäller att presentera sammansättningar såsom uppslagsord. Sammansättningar som i den äldre ordboken står i spärrad stil har räknats, trots att de egentligen inte har högre informationsvärde än dem som ordboken -86 anför i kursiv för att ge exempel på ett ords användning såsom för- eller efterled.

Ordboken från 1948 är illustrerad. Sannolikt har den flest illustrationer av de svenska ordböcker som har tagit bilder till hjälp för att förklara betydelser. Illustrationerna är sammanlagt ca 2.400, ungefär 2,5 per sida. Det har lagts ner stort arbete på bilderna, som består av både teckningar och fotografier, oftast distinkta och tydliga. Enligt förordet har bl. a. två konstnärer, R. Kempe och Ludvig Johnsson, en teckningslärarinna Elin Lund och herr Erland Budde bidragit med teckningar. Klichéer har ställts till förfogande eller inköpts från flera namngivna förlag och tryckerier. Trots det rika illustrationsmaterialet uttrycker recensenten i Nordisk tidskrift för dövstumskolan försiktigt att det måhända hade kunnat vara ännu mer omfattande.

Huvuduppslagsorden i halvfet stil är lätta att hitta. De följs av uttalsuppgift i flertalet av de fall då ordet har ett uttal som inte är givet utifrån svenskans stavningsregler. Uttalsangivelsen är grov och det vanliga alfabetet används. Apostrof sätts ut i stavelse med huvudtryck, omedelbart efter lång vokal, och efter konsonant om föregående vokal är kort. Grav eller akut accent anges inte.

Att det finns uttalsuppgifter i en ordbok för döva kan verka förbryllande, men när ordboken kom ut undervisades döva för att lära sig tala. Den mesta färdigheten tillägnade sig barnen naturligtvis via lärares och andras tal och instruktioner, men skriftliga uttalsuppgifter i en ordbok kunde vara till hjälp de gånger orden lärdes in via skrift. Traditionen att lära barnen talad svenska förstärktes under 50-talet då det tillhandahölls bättre hörapparater, men på 70-talet inträffade ett radikalt trendbrott. Undervisningen är sedan dess inriktad på att i första hand lära barnen teckenspråk, som blivit erkänt som de dövas modersmål. Skriven svenska är numera ett andra språk och talad svenska något som det läggs föga vikt vid. Redan innan teckenspråk blev förstaspråk var barnens uttal så utan nyanser att t. ex. betonad verbpartikel inte urskiljdes, inte heller en sådan företeelse som grav och akut accent. Därför är det inte konstigt att ordboken inte ger några upplysningar om sådant.

På substantiv och verb ger ordboken tillräckliga upplysningar för att man ska förstå vilka böjningsformer ordet har. Alternativa böjningar meddelas sällan. Det upplyses även hur adjektiven kompareras, men inte på ett enhetligt sätt. På t. ex. *stor* anges komparationsformerna efter uppgiften om uttal, men ibland får informationen läsas ut via språkproven (*grov*). När ändelserna är *-are*, *-ast* har det vettigt nog inte ansetts nödvändigt med vare sig direkt eller indirekt upplysning. I övrigt informeras inte explicit om morfologi eller om grammatik över huvud taget. Upplysning om ordklassstillhörighet

t.ex. ges inte, och huruvida ett verb är transitivt eller intransitivt får man utläsa indirekt ur betydelseförklaringen eller exemplifierande språkprov.

Svensk ordbok -48 uppger inte om bruklighet eller ordens stilvalör.

Svenska Akademiens ordlista nämns i förordet såsom hjälpkälla, tillsammans med SAOB, Nusvensk ordbok (A — spana), Sv. uppslagsbok, Nordisk familjebok, Norstedts uppslagsbok, Tidens lexikon, Ekborhns 60.000 främmande ord, Dalins ordbok, Auerbachs Svensk-tysk ordbok och Svensk ordlista av Lyttkens och Wulff. Sannolikt har Svenska Akademiens ordlista inspirerat till arrangemanget att i samma stycke som huvuduppslagsordet presentera ord som är etymologiskt och semantiskt samhöriga med detta, t. ex. sammansättningar, avledningar och samma ord tillhörande en annan ordklass. Ordboken -48 går i några fall ett steg längre. På t. ex. **ljus**, substantiv, står adjektivet redovisat bara såsom en tredje betydelse med översättningen "som har ljus (inte mörk) färg". Att på detta sätt t. o. m. avstå från att markera den annorlunda ordklassen medelst böjningsändelse är förvisso inte ordlistans lösning. De sammansättningar, avledningar och liknande som är underordnat redovisade står oftast i alfabetisk ordning. På **fall** t. ex. står *falla*, *fallenhet*, *fallfrukt*, *fallfärdig*, *fallgrop*, *fallrep* (med tyvärr felaktig förklaring), *fallrepstrappa*, *fallskärm*. *Notorisk* finner man underordnat på **notis**. Dessa underordnade uppslagsord har aldrig uppgift om uttal och oftast ingen böjningsuppgift, vilket i de flesta fall ju vore onödigt. Men vid exempelvis verbet *falla* upplyses om böjning. Tyvärr är de underordnade uppslagsorden ganska svåra att hitta, eftersom de är satta i en tunn, spärrad, rak stil som inte är tillräckligt iögonenfallande. Dessutom saknar de annan markering. De skulle t. ex. ha kunnat föregås av ett streck. Ibland redovisas de ganska överraskande genom att ingå i ett exemplifierande språkprov.

Betydelsebeskrivningen

Den största mödan vid utarbetandet av ordboken torde ha lagts ner på att förklara vad orden betyder. När betydelseorna redovisas uttryckligen sker det antingen genom synonym eller förklaring. Vad orden betyder framgår också via de generöst medtagna språkproven, både genom att ordets användning visas och ibland genom att hela språkprovet omskrivs eller förklaras. Att redovisa betydelse via språkprov är ibland den enda anlita-de metoden, så t. ex. på **huvud**.

Nedan jämförs hur orden **vapen**, **ärva** och **huvud** får betydelsen redovisad i Svensk ordbok -48 med hur samma ord behandlas i Svensk ordbok -86. Dessa ord har valts för att de verkar vara representativa i fråga om betydelsebeskrivningens karaktär i de bägge ordböckerna.

I Svensk ordbok -48 uppges **vapen** betyda "redskap, som begagnas till att strida med". Det första språkprovet är samtidigt en ostensiv definition: *Gevär, kanoner, sablar, spjut osv. äro vapen*. Ytterligare två språkprov ges där uttrycken *grep till vapen* ("började kriga") och *lägga ner vapnen* ("sluta att strida") används. De översätts i en inskjuten parentes.

I Svensk ordbok -86 däremot redovisas betydelsen i tre moment. Det första ger defi-

tionen ”redskap som används för att döda eller skada människa eller djur i krig, vid annan strid el. vid jakt; mest om specialkonstruerade redskap”. Detta att vapen också kan användas för att skada och att de kan begagnas vid jakt på djur har ordboken -48 missat eller underlåtit att redovisa. Ordboken -86 meddelar dessutom explicit sådana utvidgade och bildliga användningar som i *hornen är buffelns vapen* resp. *strejkvapen*. Den ger också idiomet *blanka vapen*, med två betydelser, samt förklarar uttrycken *gripa/ta till vapen*, *nedlägga vapnen*, *sträcka vapen*, *tunga vapen*, *vara under vapen*. De ytterligare momenten redovisar betydelserna ”försvarsgren” (synonym: vapenslag) resp. ”igenkänningstecken sammanfattat efter heraldiska regler”. Informationen är alltså rätt mycket rikare i ordboken -86.

Med ordet **ärva** blir jämförelsen också till nackdel för den äldre ordboken i fråga om fullständighet i redovisningen av betydelser och användningar. Men den enkla och ändå precisa definitionen i -48 av grundbetydelsen är ett drag som återkommer ofta: ”få arv (pengar l. andra ägodelar) efter en person som har dött.” Eftersom *arv* förklaras inom parentes behöver man inte slå upp detta ord, vilket hursomhelst har följande definition. ”Egendom eller annat som man erhåller efter avlidna släktingar.” Av exempel framgår därefter att såväl en soffa som vackra ögon och en god uppfostran kan vara arv. Ordboken -48 framhåller aldrig uttryckligen utvidgningar, allmänna användningar eller bildliga, vilket Svensk ordbok -86 ofta gör, men i bästa fall och inte så sällan redovisar ordboken -48 dem i exempel.

Ordboken -86 ger följande definition av grundbetydelsen hos **ärva**. ”Överta genom arv med avs. på egendom etc.” Här är det nödvändigare att ta reda på vad **arv** betyder: ”övergång av egendom av (visst) värde från avliden till efterlevande spec. som jur. term om sådan övergång genom lag.” Det må vara att ordboken -86 är precisare i sin definition, men -48 kommer enklare och med ett lättillgängligare språk fram till ungefär detsamma. Om ambitionen bakom ordboken är att erbjuda hjälp åt dem som har mindre rikt ordförråd, så är formuleringssättet i Svensk ordbok -48 många gånger exemplariskt. Den äldre ordboken är ofta tydlig också genom att undvika abstraktioner och som på **arv** framhålla den konkreta betydelsen. I rättvisans namn ska det emellertid framhållas att ordboken -86 på **arv** även tydligt redovisar utvidgningar, vilket den som sagt ofta gör och varigenom den är mycket förtjänstfull.

På **huvud** ger ordboken -48 ingen markerad betydelsebeskrivning alls, men de två första språkproven förklarar vad ett huvud är: ”På huvudet sitta ögon, öron, näsa och mun. Kon har horn på huvudet.” Dessa två exempel motsvaras i ordboken -86 av definitionen i moment 1: ”avslutande, översta eller främsta del av kroppen (på människa eller djur), där centra för de viktigaste sinnesorganen samt hjärnan är placerade ibl. betraktad som en från kroppen skild del.” (Sic!) Efter denna definition av grundbetydelsen följer tio språkprov och därefter utvidgningar, specialiseringar och idiom. I Svensk ordbok -48 redovisas utvidgningarna osv. fr. o. m. tredje kursiverade språkprovet, där det ges några uttryck med översättning: *Blotta huvudet* ”= ta av hatten”. *Bära huvudet högt* ”= vara stolt”. *Köra huvudet i väggen* ”= försöka med våld göra ngt, som är omöjligt”. etc. Även det som ordboken -86 redovisar i moment 2: ”avslutande, rund eller ngt utskjutande, del av (avlångt) föremål”, ges i ordboken -48 bara i form av språkprov: *Knappnålen har ett*

runt huvud. Svensk ordbok -48 har alltså långt ifrån den tydlighet i betydelseredovisningen som Svensk ordbok -86 uppnår genom definitioner och markerade utvidgningar, men det bör framhållas att ordboken -48 har rikhaltig exemplifiering som ganska väl illustrerar såväl centrala betydelser som utvidgningar och bildliga användningar.

För att redovisa och klargöra betydelsen tar ordboken -48 ibland motsatt betydelse till hjälp. Antonymer anges med omslutande (), två parentestecken varav det första brukar åtgöra och det andra åtgör vänster.

Encyklopedisk färgning

Svensk ordbok -48 har inte ålagt sig det stränga definitionskrav som Svensk ordbok -86 har. I stället har en strävan att förklara, hjälpa och upplysa lett till att ordboken i vissa stycken snarare blivit en encyklopedisk uppslagsbok. På **menstruation** t. ex., där synonymerna månadsrening och reglering ges, utnyttjas alla exempel för att upplysa om vad menstruation är, hur ofta den inträffar, hur länge den pågår, hur stor del av sitt liv som kvinnan får menstruation och att den upphör vid havandeskap. Svensk ordbok -86 ger definitionen "månatlig blödning från livmodern", med tillägget "då befruktning ej skett". Som exempel står två sammansättningar: menstruationscykel och -rubbning. Bra definition och kanske rimlig återhållsamhet med exempel, men man bör ha respekt för den upplysningsvilja och omsorg som ligger bakom artikeln i Svensk ordbok -48.

På **solv** tar den encyklopediska informationen fullständigt överhanden, och artikeln hör, om man bortser ifrån att den har uttals- och böjningsuppgift, helt och hållet hemma i en sakförklarande uppslagsbok: "Solv användas vid vävning. Ett solv är knutet av starkt garn, så att tre öppningar bildas. I ögat (den lilla öppningen i mitten på solvet) trädades varptråden. I de två andra, större öppningarna trädades solvkäpparna. Se väva!" En tydlig teckning bidrar till informationen vad solv är. Svensk ordbok -86 ger en kort och koncis definition, följt av tre sammansättningar såsom exempel: "var och en av en lång rad starka trådar med öglor som en varp träs igenom: *solvkäpp; solvöga; flatsolv*."

Ordbokens användbarhet idag

Man kan inte annat än hysa stor beundran och aktning för de tre dövlärare som la ner mycket tid och möda att för ringa ersättning författa Svensk ordbok 1948. Jag betvivlar att de har uppmärksammats efter förtjänst. Det är vidare djupt beklagligt att Svensk ordbok -48 inte har varit i flitigare bruk i den vanliga folk- och grundskolan. Den skulle också ha kunnat och fortfarande kunna vara till stor nytta för invandrare och andra som, liksom de döva, inte har svenska som modersmål. Dessutom bör ordboken komma i erkänt bruk när det skrivs andra ordböcker, då den nog kan inspirera. Det vore nämligen en fröjd att se en svensk ordbok som dels likt Svensk ordbok -86 är systematiskt genomarbetad så som en definitionsordbok kräver, dels likt Svensk ordbok -48 har enkla betydelseförklaringar utan onödiga abstraktioner⁴ och strävar efter att lättbegripligt och tydligt upplysa om det centralt distinkta och försöker undvika misstaget att spontanistiskt

meddela sanningar.⁵ Om ordbokens syfte är att ge förklaringar så att en språkbrukare ska kunna använda orden, så åskådliggörs och förklaras grundbetydelsen hos många vanliga ord bra bara med hjälp av exempel. Om ordboken i sådana fall ändå ger en definition är det extra viktigt att den formuleras enkelt men ändå distinkt, och att exemplen väljs omsorgsfullt. Det är stor risk att definition plus exempel annars blir förvirrande hellre än upplysande eller att arrangemanget blir helt poänglöst.⁶ Det vore dessutom önskvärt att en svensk ordbok gick ett steg längre än -86 och åtog sig att med rimliga fullständighetskrav informera en liten men studieivrig grupp som söker tydlig och systematisk upplysning om ordens syntaktiska konstruktionsmöjligheter.⁷ Den ordbok som nu är under utarbetande, Nationalencyklopedins ordbok, utlovas bli försedd med upplysningar av det slaget.

Noter

- 1 Jag tackar förre dövsläraren Börje Knutsson, Lund, för dessa upplysningar. Han har också haft vänligheten att kortfattat orientera mig om undervisning av döva för några decennier sedan och nu, och leta fram en recension i Nordisk tidskrift för dövstumskolan.
- 2 Enligt recension av signaturen I. M. i Nordisk tidskrift för dövstumskolan.
- 3 Enligt muntlig uppgift från Börje Knutsson, Lund.
- 4 Jfr dels ordboken -48 **ärva**: "få arv (pengar l. andra ägodelar) efter en person som har dött" med ordboken -86: "överta genom arv med avs. på egendom etc., dels ordboken -48:s definition av **arv**, där den konkreta betydelsen beskrivs: egendom eller annat, som man erhåller efter avlidna släktingar" med ordboken -86, som utgår från den abstrakta betydelsen: "övergång av egendom av (visst) värde från avliden till efterlevande spec. som jur. term om sådan övergång genom lag." Det bör återspeglas i en ordboksartikel att man i allmänhet i fråga om arv mer tänker på det man får än på detta att få.
- 5 Jfr definitionstillägget "mest om specialkonstruerade redskap" i ordboken -86 på **vapen**.
- 6 Ordboken -48 förklarar grundbetydelsen hos **draga** bara medelst exempel. Ordboken -86 har däremot en detaljerad definition: "påverka med en kraft som är riktad tillbaka mot kraftkällan med eller utan resultatet att det påverkade föremålet flyttas närmare". Definitionen täcker emellertid inte *dra* i uttrycket *dra ner rullgardinen* eller i *dra lakan*, med dess speciella betydelse. Dessa uttryck står ändå bland språkproven för att illustrera grundbetydelsen.
- 7 Som t. ex. Longman's eller Collins.

Litteratur

- Collins COBUILD English Language Dictionary*. 1987. Collins, London/Glasgow.
Longman's Dictionary of Contemporary English. 1978. Longman, London.
Nationalencyklopedins ordbok. Under utarbetande. Bra Böcker, Höganäs.

Carl-Erik Lundbladh

Nordisk tidskrift för dövstumskolan nr 9, 1948.

Sigurd, B. 1968: *Svenska ordböcker*. Liber, Lund.

Svensk ordbok. 1948. Svenska dövstumläraresällskapet, Lund.

Svensk ordbok. 1986. Esselte, Stockholm.

Wörterbücher. Ein internationales Handbuch zur Lexikographie. Utgiven av Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta. (Handbücher zur Sprach- und Kommunikationswissenschaft. Band 5.2) 1990. Walter de Gruyter, Berlin/New York.

Halvfasta fraser

- ett lexikografiskt problem

Erika Lyly

I artikeln diskuteras halvfasta fraser, dvs. fraser som i fråga om syntaktisk och lexikal fasthet intar en mellanställning mellan fullt produktiva konstruktioner och fasta fraser. Sådana halvfasta fraser redovisas i gängse ordböcker vanligen som tillfälliga exempel, men i den fraseologiska ordbok som för närvarande är under arbete på Svenska språknämnden kommer dessa fraser att presenteras formaliserade, i en enkel och avskalad form.

Det är viktigt att ordboksredaktören har klart för sig vilka systematiska språkliga egenskaper sådana halvfasta fraser uppvisar. Vidare bör han ha satt sig in i när en ordboksanvändare har behov av att slå upp halvfasta fraser. Frågor om urval och redigering hör nära samman med hur man betraktar och behandlar sådana fraser i det lexikografiska arbetet.

Kategorin halvfasta fraser

Vid Svenska språknämnden i Stockholm utarbetas för närvarande en fraseologisk ordbok som tills vidare går under arbetsnamnet **Svensk konstruktionsordbok**. Se Ulla Clauséns artikel "Idiom och variation" i denna skrift. **Halvfasta fraser** är en av de fraskategorier som upptar en förhållandevis stor del av det frasutbud som kommer att ingå i ordboken. Det är under arbetet med konstruktionsordboken som vi har stött på problemet med att identifiera och presentera dessa halvfasta fraser.

Med halvfasta fraser avses den fraskategori som ofta brukar gå under beteckningen "kollokationer". Men eftersom kollokationer som term har en varierande räckvidd hos olika forskare har jag valt att använda den deskriptiva termen halvfasta fraser.

Kategorin halvfasta fraser grundas på ordens kombinerbarhet, på vilka möjligheter och restriktioner som finns när man ska sammanföra olika ord. Det är inte alltid självklart vilka adjektiv och substantiv, inte heller vilka verb och substantiv som kan kombineras med varandra. Det kan också finnas en osäkerhet om vilka prepositioner som hör ihop med vilka nominalfraser. Som exempel på halvfasta fraser kan nämnas nominalfraserna *skarpt hörn*, *skarp vinkel* / *profil*, verbfraserna *bedriva forskning* / *studier* / *försöksverksamhet* och prepositionsfraserna *i vid* / *vidsträckt* / *inskränkt* / *sträng bemärkelse*.

Vilka språkliga egenskaper är utmärkande för fraser av det här slaget? Syntaktiskt är de produktiva, vilket betyder att den syntaktiska strukturen inom fraserna lätt kan om-

formas utan att de i fraserna ingående ordens betydelse förändras på något sätt. Denna syntaktiska flexibilitet gör att de halvfasta fraserna lätt bl.a. kan böjas som t.ex. frasen *skarpt hörn* i satsen *Fila bort alla skarpa hörn och kanter!* eller byggas ut med attributiva bestämningar som t.ex. frasen *bedriva försöksverksamhet* i satsen *Det finns resurser att bedriva en ingående försöksverksamhet*. Lexikalt uppvisar de halvfasta fraserna däremot en viss fasthet, vilket yttrar sig i att det finns vissa restriktioner för vilka ord som kan kombineras med varandra. I t.ex. prepositionsfrasen *i vid bemärkelse* står *vid*, men inte *stor* som bestämning till *bemärkelse*. Det är denna lexikala egenskap att inte vara helt förutsägbara som gjort att de halvfasta fraserna fått samma dignitet i konstruktionsordboken som övriga mer etablerade fraskategorier.

Presentationen av halvfasta fraser och andra fraser i konstruktionsordboken

Målet för arbetet med konstruktionsordboken har hela tiden varit att lyfta upp det strukturella, vilket avspeglar sig i det sätt på vilket det fraseologiska materialet presenteras. För det första ges fraserna, vilket framgår av fig. 1, i **hierarkisk ordning**, från fria, produktiva fraskonstruktioner till fasta lexikaliserade fraser, idiom och andra funktionella uttryckssätt. För det andra är de olika **nivåerna i framställningen** klart åtskilda. Det betyder att fraser, exempel och kommentarer presenteras med olika typografi och i olika språklig form. (Se artikeln **bekänna**, fig. 2). Alla fraser, oavsett typ av fras, presenteras i en formaliserad och avskalad form, som s.k. formaliseringar och utmärks typografiskt med halvfet stil. Dessa formaliseringar är visuella ingångar till en liten text-helhet i vilken oftast exempel och olika slag av kommentarer kan ingå.

KONSTRUKTIONSORDBOKEN

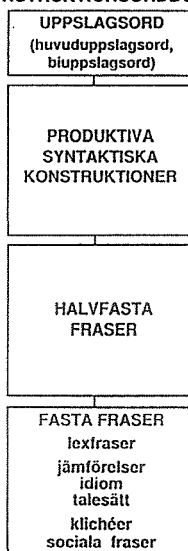


Fig. 1

bekänna verb 1 (erkänna, medge) *nqn*
bekänner har han bekänt?; *bekänna ngt*
han har bekänt sin skuld; han bekände
att han var skyldig till mordet på den
gamla kvinnan; åld. *bekänna sin kärlek*
han bekände sin kärlek för henne; i
kortspel; [*inte*] *bekänna färg* se *färg*,
bekänna kort se *kort*; ursäktande *jag*
måste bekänna att jag måste ~ att jag
har ätit upp kakorna; *jag måste* ~ att
jag aldrig läst en deckare
2 (förklara sig tro på ngt) *bekänna*
ngt ~ sin kristna tro; hon bekände sin
tro på psykologin
bekänna sig refl. 1 (medge, erkänna)
skriftspr. *bekänna sig skyldig till ngt*
han bekände sig skyldig till mordet
på den gamla kvinnan 2 (förklara sig
tro på något) *bekänna sig till ngt* i
sin ungdom bekände han sig till
kommunismen

Fig. 2

Artikeln *bekänna* är ett exempel på hur ordartiklarna kommer att se ut i den färdiga ord-boken. Avgörande för var i artiklarna de enskilda fraserna placeras (fig. 1) är, som fram-gått tidigare, deras grad av syntaktisk fasthet. De halvfasta fraserna, som ju är syntak-tiskt produktiva, är placerade i början av artikeln, men på andra plats, eftersom de sam-tidigt är lexikalt fasta. Ordartiklarna inleds med **de produktiva syntaktiska konstruk-tionerna**. Dessa anger vilka syntaktiska mönster som är möjliga vid ett aktuellt uppslags-ord. Man kan också säga att de produktiva syntaktiska konstruktionerna visar på ett ur-val tänkbara valenser. Ordboksanvändaren kan känna igen den här typen av konstruk-tioner på att alla ord i formaliseringarna är formord förutom uppslagsordet, som av na-turliga skäl alltid är ett innehållsord. Jämför *ngn bekänner, bekänna ngt*.

De **fasta fraserna** utgör en mycket stor och på många sätt heterogen grupp av fra-ser. Dessa fraser är syntaktiskt fasta även om vissa undantag kan förekomma. Se Ulla Clauséns artikel "Idiom och variation". Den grupp av fasta fraser som på många sätt lik-nar de halvfasta fraserna är de s.k. **lexfraserna**. Lexfraserna och de halvfasta fraserna är lika varandra i väsentliga avseenden. De hör båda till lexikonet och de är lika till for-men, dvs. de uppvisar exakt samma syntaktiska strukturer. För att skilja de halvfasta fra-serna från lexfraserna har vi använt de kriterier för lexfraser som Jan Anward och Per Linell ställde upp i sin klassiska artikel "Om lexikaliserade fraser" i Nysvenska studier 1976 (s. 109). Dessa kriterier går sammanfattningsvis ut på att lexfraserna är syntaktiskt oföränderliga och semantiskt ogenomskinliga. Exempel på typiska lexfraser är *bekänna färg* och *bekänna kort*. Dessa är spelstermer som också har en metaforisk användning och då betyder 'ypa sina avsikter' eller 'tala rent ut'. Dessa fraser är syntaktiskt fasta. I sin metaforiska användning är de semantiskt ogenomskinliga. Fastheten markeras också av den sammanhållande betoningen. Hela frasen bildar en helhet som refererar till någon-ting annat än vad de enskilda orden antyder. Frasen *bekänna sin kärlek* är däremot var-ken syntaktiskt fast eller semantiskt ogenomskinlig och klassificeras som halvfast fras.

Gränserna mellan fraskategorierna är inte skarpa. Det finns fraser som representerar ett slags övergångsformer mellan lexfraser och halvfasta fraser. Verbfraserna *spela kort / piano / tennis* är t.ex. syntaktiskt fasta. Nominalen i fraserna kan varken böjas eller byggas ut med bestämmningar. Dessa fraser har m.a.o. syntaktiska egenskaper som är ut-märkande för lexfraserna. Men de är semantiskt ogenomskinliga, vilket däremot är en av de egenskaper som karaktäriserar de halvfasta fraserna. Exempelfraserna för inte fram någon annan betydelse än den som de enskilda orden antyder. Vid fraser av det här sla-get med egenskaper som syntaktiskt ger frasen en klassificering och semantiskt en annan har de syntaktiska kriterierna fått vara avgörande för val av fraskategori. Verbfraserna ovan har alltså klassificerats som lexfraser och inte som halvfasta fraser.

Ett annat problem av rent syntaktisk natur är att fastställa var frasgränserna går, dvs. att avgöra var fraserna börjar och var de slutar. Detta problem kommer tydligt fram vid utformningen av formaliseringarna för de halvfasta fraserna. Dessa uppträder ofta i en annan form än i den rena, avskalade formen som de syntaktiskt fasta lexfraserna gör. Problemet aktualiseras framför allt när det gäller att avgöra om en halvfast fras ska pre-senteras som verbfras eller inte. Det finns många halvfasta fraser som i kontext alltid står

som verbfraser. Är t.ex. prepositionsfrasen *under bekvämlighetsflagg* en fullständig fras eller borde den redovisas som *segla / gå under bekvämlighetsflagg*? I det här fallet går det att konstruera satser där frasen står som ren prepositionsfras utan att samtidigt ingå i en verbfras. Ett sådant fall förekommer t.ex. i satsen *X har en hel flotta under bekvämlighetsflagg*. Frasen har alltså redovisats som prepositionsfras.

Halvfasta fraser i ett användarperspektiv

Att vi lyft fram de halvfasta fraserna och låtit dem framstå som en egen fraskategori beror som tidigare nämnts på att dessa fraser lexikalt inte är helt förutsägbara. Det heter t.ex. *fälla ett yttrande*, men inte **uttala ett yttrande* och *hårt väder*, men inte **strängt väder*. Varför orden går att kombinera i det ena fallet men inte i det andra finns det inget självklart svar på. Vissa ordkombinationer har av olika anledningar konventionaliserats och blivit allmänt godtagbar norm medan andra, som i och för sig skulle kunna vara logiskt tänkbara, framstår som oacceptabla. Det är därför inte självklart att man som skribent eller översättare alltid kommer på vilka ord som är lämpliga att kombinera med varandra. Det är då man kan ha behov av fraseologiska ordböcker där detta går att slå upp i.

För ordboksredaktören är det viktigt att sätta sig in i när en språkbrukare kan behöva slå upp halvfasta fraser. Vid uppläggningsen av konstruktionsordboken har vi fram för allt beaktat följande två fall. För det första ska ordboksanvändaren kunna slå upp ett representativt urval lexikala kombinationer för det ord som han vill bygga ut till en fras. För det andra ska användaren kunna få information om i vilka semantiska eller begreppsmässiga sammanhang ett uppslagsord kan förekomma. Dessa båda sökningar kräver olika redaktionella hänsyn. I det första fallet söker man i regel på det ord i frasen som är semantiskt tungt eller markerat. Detta står normalt finalt. Det ord som är semantiskt tungt är frasens huvudord och samtidigt det naturliga uppslagsordet för den som letar efter lämpliga kombinationsmöjligheter. I det andra fallet använder man det ord som frasen inleds med som uppslagsord. De inledande orden tenderar att ha en mera generell betydelse än de ord som står finalt. De behöver inte för den skull vara semantiskt lätta eller omarkerade, men de har en mindre specifik betydelse än de finala huvudorden. I det här fallet är det av vikt att de halvfasta fraserna är sorterade så att man genom att studera dem kan få en uppfattning om i vilka semantiska sammanhang uppslagsordet kan användas.

Det semantiskt tunga ordet i en halvfast fras är i de flesta fall ett substantiv, oavsett frasens grammatiska struktur. Man kan alltså räkna med att substantiven är de ord som används som sökord när man letar efter lämpliga kombinationsmöjligheter. En av de frågor som en ordboksanvändare kan behöva få svar på är t.ex. frågan om vilka verb som går att kombinera med ett bestämt substantiv för att sedan bland de givna alternativen snabbt kunna välja det verb som passar bäst i den aktuella kontexten. Under uppslagsorden *begäran* och *skilsmässa* kan följande halvfasta fraser sammanföras i formaliseringarna *avslå / bifalla / bevilja en begäran*, *söka / begära / ta ut skilsmässa* och *få / gå med på*

/ *medge skilsmässa*. De verb som ges inom en formalisering utgör här ett urval av existerande möjligheter och är enbart exempel på användbara alternativ. Som framgår av exemplen ovan hör verben till samma semantiska typ. Om verben inte gör det och t.ex. har olika kasusroller, sorteras de inte inom samma formalisering utan delas upp på flera. Jämför formaliseringarna till *skilsmässa*.

Det är vanligt att många halvfasta fraser inleds av verb, men de kan naturligtvis också inledas av attributiva bestämningar och då gärna av ett adjektiv. Dessa inledande ord har som redan nämnts en mindre specifik betydelse än de finala orden. Detta gör att de kan ha ett tämligen brett betydelseregister och att de kan användas i olika innehållsmässiga och stilistiska sammanhang. T.ex. adjektivet *skum* ingår i vårt ordboksmanus i följande halvfasta fraser *skum belysning*, *skum typ*, *skum verksamhet*, *skumma affärer / förehavanden / syften*, *skumma rykten*. Betydelsenyanserna i exempelfraserna är 'lusfattig', 'mystisk', 'tvivelaktiga' och 'svårtydda / betänkliga'. Vilka betydelsenyanser adjektivet *skum* har bestäms av det substantiv som det är attribut till. Många betydelser illustreras bäst genom att språkbrukaren får se hur orden kombineras. Även om halvfasta fraser är semantiskt genomskinliga och därför inte kräver några betydelseförklaringar kan det ändå vara befogat att ge sådana när det ena ledet har en mindre specifik betydelse.

Under arbetet med konstruktionsordboken har vi kommit fram till att halvfasta fraser är en viktig kategori om än icke helt problemfri. I många ordböcker har de halvfasta fraserna inte varit så väl representerade kanske därför att de ansetts vara självklara till sin betydelse. Eftersom den nya konstruktionsordboken i huvudsak är tänkt att vara en ordbok för den aktiva skribenten behövs ett generöst urval av halvfasta fraser.

Litteratur

- Anward, J. & Linell, P. (1976) *Om lexikaliserade fraser*. Nysvenska studier, Årg. 55-56, s. 77-119.
- Benson, M. (1989) *The Structure of the Collocational Dictionary*, International Journal of Lexicography, Vol. 2, Nr. 1, s. 1-14.
- Boje, F. & Braasch, A. (1992) *Hvad får man skudt i skoene? Flerordsenheder i aktive ordbøger for mennesker og maskiner*. Nordiske studier i leksikografi, Skrifter utgitt av Nordisk forening for leksikografi, Skrift nr. 1, Oslo, s. 483-500.
- Burger, H. (1982) *Handbuch der Phraseologie / von Harald Burger, Annelies Buhofer u. Ambros Sialm. Walter de Gruyter, Berlin*.
- Carter, R. (1987) *Vocabulary*. Allen & Unwin, London.
- Dixon, R.M.W. (1991) *A New Approach to English Grammar on Semantic Principles*. Clarendon Press, Oxford.
- Gellerstam, M. & Malmgren, S.-G. & Toporowska-Gronostaj, M. (1992) *Inlärningslexikon*. Nordiske studier i leksikografi, Skrifter utgitt av Nordisk forening for leksikografi, Skrift nr. 1, Oslo, s. 347-356.
- Gläser, R. (1986) *Phraseologie der englischen Sprache*, VEB Verlag Enzyklopädie Leip-

Erika Lyly

zig.

Hausmann, F.J. (1985) *Kollokationen im Deutschen Wörterbuch*, Lexikographie und Grammatik, Hrsg von Bergenholtz & Mazdan, Tübingen: Niemeyer [Lexicographica Series Maior 3].

Moon, R. (1992) *'There is reason in the roasting of eggs': a consideration of fixed expressions in native-speaker dictionaries*. EURALEX '92 Proceedings II, Tampere, s. 492-502.

Svensén, B. (1987) *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. Stockholm.

van der Wouden, T. (1992) *'Prologomena to a Multilingual Description of Collocations'*. EURALEX '92 Proceedings II, Tampere, s. 449-456.

Sammanställningsmorfologi och lexikografi

Sven-Göran Malmgren

I Illustrerad svensk ordbok, den svenska definitionsordbok som kom ut på 1950-talet, finner man bland sammansättningarna under **röd** bl.a. en kort artikel **rödbena** (en fågel) och strax därefter **rödbeneägg**, med förklaringen 'ägg av rödbena'. Vid första påseendet kan den sista artikeln verka totalt överflödig. Man kan omöjligt hävda att sammansättningen **rödbeneägg** är lexikaliserad, och definitionen utmanar nästan löjet. Men det finns faktiskt ett bra skäl till att ordet **rödbeneägg** är med i ordboken. Det är inte alls självklart vilket utseende ordet **rödbena** får som förled i sammansättningar. Vad artikeln upplyser om är i själva verket just att den rätta formen är **rödbene-** och inte t.ex. **rödbens-** (jfr *pepparkaka - pepparkaksdeg*). Den i ordboken upptagna sammansättningen ska alltså ses som representant för alla möjliga sammansättningar som utgår från **rödbena**. (Definitionen kunde dock med fördel ha utelämnats.)

Om alltså artikeln **rödbeneägg** faktiskt innehåller en värdefull upplysning, så kan man likväl fråga sig om ordbokens sätt att ge denna upplysning är det bästa. Man kan också fråga sig hur konsekvent uppgifter om sammansättningsfogar hos ord som själva är sammansatta ges i moderna svenska ordböcker (senare än Illustrerad svensk ordbok). I den här uppsatsen görs ett försök att besvara båda frågorna, och dessutom diskuteras en hittills knappast prövad möjlighet att ge information om sammansättningsmorfologin. Slutligen berörs möjligheterna att snabbt samla in nödvändiga data från korpusar och lexikaliska databaser.

Vi kan nöja oss med att se på elfte upplagan av Svenska Akademiens ordlista (SAOL, 1986), eftersom det är den enda moderna svenska ordbok som ger nämnvärd information av det aktuella slaget. I andra ordböcker, t.ex. Svensk ordbok (1986), tas många sammansättningar upp, men så gott som bara sammansättningar som är semantiskt intressanta. Sammansättningarna av typen **rödbeneägg** torde vara mycket lätt räknade i Svensk ordbok. Vi väljer vidare att koncentrera oss på substantiv av samma typ som **rödbena**, dvs. substantiviska sammansättningar som slutar på **-a**. Hur dessa ord förändras när de själva ingår som förled i dubbelt sammansatta ord av typen ABC är synnerligen oförutsägbart, som Hellberg (1978) och framför allt Wennerberg (1962) visar. Jfr t.ex. *ficklampa - ficklampsbatteri*, *folkvisa - folkvisetradition*, *högmässa - högmässobesökare*, *vattenskada - vattenskidåkare*. Tilläggas kan, att problemen är likartade vid sammansättningar vars efterled slutar på **-e**. När efterleden slutar på konsonant är det betydligt lättare att förutsäga sammansättningsmorfologin; vi tar emellertid upp ett komplicerat fall i slutet av uppsatsen.

Vi ser först på några substantiv som liksom **rödbena** betecknar fåglar, närmare be-

stämt alla sammansättningar med *-kråka*, *-lärka*, *-skata* eller *-ärsla* som efterled som tas upp i SAOL. (Det finns en baklängessortering av ordboken.) Fördelen med att studera denna "tematiska" grupp är att den troligen inte genererar många lexikaliserade sammansättningar. Om alltså en dubbel sammansättning (t.ex. *sädesärleägg*) kommit med i SAOL, kan man räkna med att det är av samma skäl som att *rödbeneägg* togs upp i Illustrerad svensk ordbok, och det är det som är det intressanta i det här sammanhanget.

I tabell 1 redovisas resultatet av den lilla undersökningen:

Tabell 1. Information om sammansatta ord som förleder i SAOL

| Ord | Sammansättning med ordet som förled finns/finns inte (+/-) i SAOL (1986) |
|--------------|--|
| blåkråka | - |
| nötkråka | - |
| spillkråka | + (spillkråksbo) |
| piplärka | - |
| skärpiplärka | - |
| sånglärka | - |
| tofslärka | - |
| trädlärka | - |
| snöskata | - |
| strandskata | - |
| törnskata | - |
| gulärsla | - |
| sädesärsla | + (sädesärleägg) |

Vi ser att en språkbrukare som är osäker på hur man bildar en sammansättning med *sånglärka*, *tofslärka* etc. eller med *snöskata*, *strandskata* eller *törnskata* inte får någon vägledning av ordboken. Vill han bilda sammansättningar med *nötkråka* eller *gulärsla* får han inte heller någon **direkt** upplysning, men med lite tur och fantasi kan han hitta fram till *spillkråka* resp. *sädesärsla* och på grundval av sammansättningarna där gissa att det bör heta t.ex. *nötkråksbo* och *gulärleägg*. Men endast när det gäller *spillkråka* och *sädesärsla* får han explicit vägledning.

Nu finns det givetvis en gräns för hur många ord SAOL kan ta upp, och meningen är

ingalunda att hävda att alla de saknade dubbla sammansättningarna borde varit med. Men möjligen hade det varit bra om ett av de substantiv – och i så fall det man bedömer som det vanligaste – som slutar på *-lärka* och ett av dem som slutar på *-skata* hade försetts med en sammansättning, så att en användare åtminstone hade möjlighet att få indirekt vägledning beträffande de övriga orden med dessa efterleder. (En annan möjlighet att ge denna vägledning diskuteras i slutet av uppsatsen.)

Naturligtvis kan man säga att det inte är någon katastrof om en språkbrukare själv måste bestämma sig för om det ska heta *sånglärkebo*, *sånglärksbo*, *sånglärkbo* eller något annat. Och det ska medges att problemet inte hör till de mest centrala i språket. Men frågor föranledda av osäkerhet om förledens form i sammansättningar tycks höra till de vanligaste bland de frågor som ställs till Svenska språknämnden (Ulla Clausén, muntl. komm.). Det finns också en liten paradox i det här sammanhanget. Å ena sidan har man sällan anledning att bilda dubbla sammansättningar, och därför kan man säga att problemet med fogens utseende inte är så viktigt, men å andra sidan medför just de dubbla sammansättningarnas sällsynthet att språkkänslan är osäker, och därför kan behovet av vägledning vara särskilt stort när det någon gång finns behov av en sådan sammansättning.

För att illustrera ett par andra problem ser vi på ytterligare ett par grupper av ord i SAOL:s elfte upplaga. Först sammansättningarna med *-penna* som efterled. De är 28 till antalet. Endast vid *reservoarpenna* ges en dubbel sammansättning, *reservoarpenn(e)bläck*. Ser man nu efter i den närmast föregående upplagan av SAOL (1973), finner man utom *reservoarpenn(e)bläck* även det dubbelt sammansatta ordet *stålpennsfabrik*. Det fanns säkert mycket goda skäl att stryka detta ord i elfte upplagan av SAOL. Men strykningen har faktiskt en konsekvens som möjligen inte var avsedd: med den försvinner den enda indikationen i SAOL på att sammansättningar med efterleden *-penna* kan få fogelementet *-s-* när de ingår som förled i dubbelt sammansatta ord. En ambitiös ordboks-användare, eller en konstruktör av lemmatiseringsprogram eller program för stavningskontroll – beväpnad med baklängesordbok – som förlitar sig på SAOL måste dra slutsatsen att mönstret *reservoarpenn(e)bläck* numera, i motsats till för 20 år sedan, är genomgående. Kanske är det också så, kanske inte. Poängen här är bara att strykningen av ett helt perifert ord faktiskt kan innebära ett rätt kontroversiellt implicit påstående.

De sista exemplen vi tar upp är sammansättningarna med efterleden *-stuga*. De är 47 till antalet i SAOL (1986), och vid 6 av dem ges dubbelt sammansatta ord. Dessa är *backstugusittare*, *förstu(gu)trappa* (m.fl.), *rådstu(gu)rätt*, *småstugebebyggelse*, *sommarstugebebyggelse* och *sportstugeområde*. Om man bortser från variationen vid *förstuga* och *råd-stuga* verkar alltså *-stuge-* och *-stugu-* vara de två möjligheter som finns. Nu kan man emellertid observera att *-stugu-* i SAOL endast förekommer i mer eller mindre ålderdomliga ord, och troligen används denna form sällan i nybildade sammansättningar. Kanske vore en explicit upplysning om detta av värde, även om det redan nu antyds av SAOL:s urval av dubbla sammansättningar.

På vilket sätt skulle dessa upplysningar om sammansättningsmorfologi kunna ges? Om

vi stannar vid det sista exemplet, kunde man kanske tänka sig något i stil med följande (t.o.m. **-värme** är det den autentiska artikeln i SAOL):

stug|a -an -or s. stug|by -dörr [...] -värme; -stuge-, (förr) -stugu-

Informationen om hur *tvättstuga*, *fritidsstuga* etc. bör se ut som förled i dubbla sammansättningar skulle alltså samlas under det enkla ordet *stuga*. En användare, som söker råd när han ska bilda en sammansättning av *tvättstuga* och *tid*, behöver veta att han ska gå från artikeln **tvättstuga** till artikeln **stuga**, när han inte får vägledning på det första stället. Den upplysningen skulle han givetvis få i ordbokens inledning. När han väl hittat fram till artikeln **stuga** får han en direkt rekommendation att använda formen *tvättstugetid*.

På motsvarande sätt skulle man kunna ge den information som nu saknas vid de övriga orden vi diskuterat. Man skulle då också tvingas ta ställning till om möjligheten *-penns-* eventuellt bör anges vid sidan av *-penn-* och *-penne-* i artikeln **penna**.

Den föreslagna modellen fungerar emellertid tillfredsställande bara under förutsättning att alla sammansättningar med en viss efterled följer samma mönster som förleder. När detta inte är fallet, vilket är långtifrån ovanligt, får det anges på något sätt. Vi kan ta sammansättningarna med *-vatten* som efterled som exempel. SAOL ger enbart *regnvattens-* (*regnvattenstunna*) som förledsform av *regnvatten*, enbart *grundvatten-* av *grundvatten* (*grundvattennivå*), men både *lågvattnens-* och *lågvattnen* (*lågvattnen(s)märke*) av *lågvattnen*. Av artikeln **vatten**, där möjligheterna *-vatten-* och *-vattens-* enligt den föreslagna modellen skulle anges, skulle det då behöva framgå att valet inte alltid är fritt utan att användaren bör gå vidare till de enskilda sammansättningarna med *-vatten* som efterled. Men finner han vid en sådan sammansättning inget exempel på ett dubbelt sammansatt ord, som t.ex. vid *badvatten* i den aktuella upplagan av SAOL, bör han kunna dra slutsatsen att båda varianterna är möjliga just vid detta ord.

Den information som här har föreslagits kan mycket väl ges också i definitionsordböcker. Kanske passar den t.o.m. bättre där, om inte annat så av typografiska skäl. (Definitionsordböcker brukar ha något mindre kompakt typografi än SAOL.)

Återstår frågan hur man bäst får fram det empiriska underlaget för de uppgifter om sammansättningsmorfologi som här diskuterats. För det första bör man kunna söka på strängar inne i ord i en lexikalisk databas. Den möjligheten finns redan nu vid Institutionen för svenska språket vid Göteborgs universitet (där arbetet på tolfte upplagan av SAOL och på en utvidgning av Svensk ordbok pågår). Den maskinläsbara ordlista man kan söka på detta sätt i innehåller mer än 150 000 ord, huvudsakligen från SAOL. Om man t.ex. beställer fram alla ord som innehåller strängen *stug* föregående och följd av minst två bokstäver får man efter några sekunder följande lista på 15 ord på terminalen:

Fig. 1. Ord med strängen *stug* föregången och följd av minst två bokstäver i en lexikalisk databas på 150 000 ord

Söksträng: Maxant

Lemman i L

| Grundform | Neo | Hv | Lx | Ordklass |
|-----------------------|-----|----|----|----------|
| backstugusittare | 1 | | 1 | subst. |
| förstugubro | 1 | # | | subst. |
| förstugukvist | 1 | # | | subst. |
| förstugutrappa | - | | | subst. |
| ristuggare | - | | | subst. |
| rådstugurätt | 1 | a | 1 | subst. |
| rådstugurättspredikan | - | | | subst. |
| småstugebebyggelse | - | | | subst. |
| snustuggande | - | | | adj. |
| snustuggande | - | | | subst. |
| snustuggning | - | | | subst. |
| sportstugeområde | - | | | subst. |
| testuggare | 1 | | 1 | subst. |
| tobakstuggning | - | | | subst. |

Vi får alltså fram de ord vi tidigare diskuterat och dessutom en del ord som vi inte är intresserade av. Det är uppenbart att man vill kunna söka på morfemet *stug*. Detta har ännu inte utförts i praktiken men det är i princip möjligt – morfemgränserna är markerade i de flesta av de 150000 orden.

Men för att kunna ge tillförlitliga uppgifter om sammansättningsmorfologi behöver man i många fall ha tillgång till data från en **korpus**. Idealet vore att från en korpus på exempelvis 10 miljoner ord snabbt kunna få fram grundformerna av alla ord som innehåller ett visst morfem, sorterade i första hand efter varianterna av morfemet (eller av morfemet plus den tillhörande sammansättningsfogen) och med frekvensuppgifter. För detta behöver man, utom markeringar av morfemgränser i orden i den lexikaliska databasen, ett system för automatisk lemmatisering. Vi vet att det bästa svenska lemmatiseringssystemet ger rätt analys av mer än 99% av orden i en allmänspråklig korpus (Karlsson 1992, om det s.k. SWETWOL-systemet). Många fall av homografi som normalt väl-

lar problem är harmlösa om man i första hand söker efter morfem **inne** i ord. Att från en stor korpus snabbt få fram exempelvis alla dubbelt sammansatta ord där förleden slutar på *-stugu-* eller *-stuge-* ligger därför inom möjligheternas gräns.

Ett litet problem återstår dock. Ett lemmatiseringssystem måste bygga på existerande ordböcker, t.ex. SAOL. Antag att lemmatiseringssystemets lexikon i likhet med SAOL inte räknar med möjligheten *-penns-*. Om ett ord som *kulpennsfodral* förekommer i korpusen kommer det då att höra till den knappa procent ord som lemmatiseringssystemet inte klarar att analysera. Den givna slutsatsen är att dessa ord är mycket viktiga att gå igenom eftersom de kan indikera förändringar i sammansättningsmorfologin.

Litteratur

Illustrerad svensk ordbok. 1956. Stockholm.

Hellberg, Staffan 1978. *The Morphology of Present-Day Swedish*. Stockholm.

Karlsson, Fred 1992. SWETWOL: *A comprehensive Morphological Analyser for Swedish*. *Nordic Journal of Linguistics*, 15, s. 1-45.

SAOL=*Svenska Akademiens ordlista*. 10, 11. uppl. 1973, 1986 Stockholm.

Svensk ordbok. 1986. Stockholm.

Wennerberg, John. 1962. *De dubbelt sammansatta orden i svenskan*. *Nysvenska studier*, 41, s. 5-54.

Översättarhjälp i en ordbok

Nina Martola

1. Inledning

Finska språket skiljer sig strukturellt mycket starkt från svenskan. Det hör till den finsk-ugriska språkfamiljen som inte ens är besläktad med de indo-europeiska språken. Finskan är ett syntetiskt språk och använder sig av ordbildning i många fall där svenskan utnyttjar syntax.

Det här har en påtaglig praktisk konsekvens för lexikografen som arbetar med en ordbok från finska till svenska. Han saknar helt enkelt ord, svenska ord. Det traditionella sättet att lösa det problemet brukar vara att använda sig av nakna exempel, och den möjligheten utnyttjar också vi, men det finns vissa frekventa ordtyper som i regel kräver långa exempel för att både svenskan och finskan skall bli naturlig och som representerar en problematik som snarare kan sägas vara översättningsteknisk än lexikografisk. Sådana fall planerar vi att behandla i samlande översättningsmodeller i stället för att mer eller mindre otillfredsställande försöka åskådliggöra samma problematik under vart och ett uppslagsord särskilt.

Det finns också ett annat skäl till att vi stannat för den här lösningen. Med tanke på hur mycket som översätts i Finland är inställningen till översättningsarbetet på sina håll tämligen oprofessionell och man träffar tyvärr på en hel del undermåliga översättningar. Till undermåligheten bidrar i hög grad brott mot vissa grundläggande regler för svensk satskonstruktion. Hurdana brott det rör sig om kommer att framgå i det följande.

2. Syntaktisk översättarhjälp

I denna artikel skall jag presentera fyra olika ordtyper som vi har tänkt oss att behandla i modellartiklar. De tre första typerna är alla substantiveringar, den fjärde är kausativa verb.

2.1. Verbalsubstantiv

För att åskådliggöra problematiken kring verbalsubstantiven börjar jag med ett autentiskt exempel, hämtat ur en bok om Högholmens Zoo i Helsingfors. (Upplagan drogs in på grund av den dåliga översättningen och gavs sedan ut i omarbetat skick.)

Ex. Snöleopardernas framgångsrika fostrande har gett Högholmen internationellt rykte.

Nina Martola

För att analysera satsen strukturellt byter vi ut *fostrande* till det i sammanhanget korrekta *uppfödning* och får då ett typiskt exempel på det som i Finland brukar kallas "översättningska":

Ex. Snöleopardenas framgångsrika uppfödning har gett Högholmen internationellt rykte.

För en person vars svenska språkkänsla är intakt verkar en sådan sats ytterst märklig, men om man så att säga kontrollerar frasen mot finskans syntaktiska regler är den inte anmärkningsvärd på något vis.

Om man löser upp nominalfrasen till en sats som representerar tankeinnehållet (TI) i den ser man vari felet består.

Ex. leopardien kasvatus

TI: leopardeja kasvatetaan 'leoparder föds upp'

TI "någon föder upp leoparder"

leopardernas uppfödning

TI "leoparderna föder upp något"

Felkällan är alltså att reglerna för användningen av genitiv är olika i finska och svenska. På finska kan genitiv användas för att uttrycka en mängd olika relationer. Noteras kan också att *leopardeja* är objekt i den finska satsen *leopardeja kasvatetaan* (Ikola 1986: 149 ff), inte subjekt som i motsvarande svenska passivsats.

I svenskan är genitiven i nominalfraser av den här typen i regel subjektiv, d.v.s. ordet i genitiv blir subjekt om nominalfrasen löses upp till en aktiv sats som representerar tankeinnehållet. (Reuter 1993: 21ff).

Ex. professorns föreläsning

TI "professorn föreläser"

Objektiv genitiv kan förekomma, men är inte särskilt vanligt. Vissa verbalsubstantiv, som t.ex. *uppläggnings*, kan ta antingen subjektiv eller objektiv genitiv.

Ex. föreläsningens uppläggnings

TI "någon lägger upp föreläsningen"

föredragshållarens uppläggnings

TI "föredragshållaren lägger upp något"

Eftersom *uppfödning* i leopardexemplet inte medger objektiv genitiv — leoparder kan inte syssla med uppfödning — blir exemplet konstigt för svenska öron.

Genitiv kan på finska som sagt användas för att uttrycka en mängd olika relationer, och förmodligen är det det som gör att finskan också tillåter flera genitivbestämningar efter varandra:

- Ex. eläintarhan leopardien kasvatus
- *djurparkens leoparders uppfödning
 djurparkens uppfödning av leoparder
- Helsingin eläintarhan leopardien kasvatus
- *Helsingfors' djurparks leoparders uppfödning
 Helsingfors djurparks leoparduppfödning
- Helsingin eläintarhan leopardien kasvatuksen onnistuminen
- *Helsingfors' djurparks leoparders uppfödningens framgångsrikhet
 "att Helsingfors djurpark så framgångsrikt föder upp leoparder"

Finskan kan bilda verbalsubstantiv till nästan vilket verb som helst. Verbalsubstantivet får bilda kärnan i en nominalfras som i sin tur kan fungera som satsdel och få bestämningar. Språket förfogar därmed över ett effektivt medel till komprimering och invävning. Följande exempel är en typisk tidningsrubrik, hämtad ur Helsingin Sanomat 10 maj 1993:

- Ex. Iijoen pinnan nousu hidastui
- *Ijo älvs ytas stigning förlångsammades

Svenskan har i stället möjlighet att låta bisatser och infinitiv få satsdelsfunktion på ett sätt som finskan saknar. Finskans infinitiver kan visserligen också utgöra satsdelar, men de är substantiviska till sin karaktär (Ikola 1986: 60 ff.) och satskonstruktionen blir därför en annan. I exemplet ovan kunde man tänka sig att byta ut *nousu* mot infinitivformen *nouseminen*.

Den finska genitivbestämningen blir lätt mångtydig, medan svenskan med hjälp av sina prepositioner kan uttrycka exakta relationer mellan orden.

Många gånger väljer svenskan ett helt annat sätt att se på saken. Rubriken om Ijo älv kunde tänkas lyda *Ijo älv stiger långsammare* i en svensk tidning.

Den som är bunden av den finska förlagan glömmar svenskans möjligheter och det är dem vi vill försöka lyfta fram med hjälp av våra syntaktiska modeller.

Vi skall se på ett annat verbalsubstantiv, nämligen *lento*, bildat till verbet *lentää* 'flyga'.

Motsvarande verbalsubstantiv på svenska är *flygning* och kanske *flygande* i den mån *lento* kan få en durativ aspekt. För vissa typers föremål motsvaras *lento* av *luftfärd* och *bana* och om fåglar används används i en viss innebörd *flykt*. De svenska ekvivalenterna har alltså mer specialicerade betydelser. Ingen av dem kan användas i en allmän betydelse 'att röra sig genom luften' oavsett vad det är som rör sig, vilket det finska ordet alltså kan.

| | | |
|-----|----------------------|--------------------------------|
| Ex. | purjekoneen lento | *segelflygplanets flygning |
| | leijän lento | *drakens flygning |
| | kanan lento yöpuulle | *hönans flygning upp på pinnen |
| | auton lento pellolle | *bilens flygning ut på åkern |

Den här typens nominalfraser är naturliga enheter i finskan och förekommer frekvent som exempel i finska ordböcker (de två första exemplen är hämtade direkt ur Suomen kielen perussanakirja). På svenska saknas ofta en motsvarande naturlig fras och den finska nominalfrasen måste sättas in i ett sammanhang för att exemplet skall fungera på svenska.

I en artikel som *lento* ges de specialiserade ekvivalenterna i själva artikeln, medan beskrivningen för den allmänna betydelsen 'att röra sig genom luften' ersätts med en hänvisning till en modell för översättning av verbalsubstantiven.

När det gäller verbalsubstantiven är problemet inte i första hand total avsaknad av ett motsvarande ord, utan problemet är framför allt den stora skillnaden mellan språken vad gäller frekvens och användningsmöjligheter.

2.2. Egenskapssubstantiv

För nästa typ av ord är problemet däremot rätt ofta total avsaknad av ett motsvarande svenskt ord. Finskan har nästan obegränsade möjligheter att bilda egenskapssubstantiv, d.v.s. substantiv bildade till adjektiv eller particip. Den typiska ändelsen är *-uus/-yys*.

På svenska har vi en motsvarighet i avledningsändelsen *-het* men den är inte alls lika produktiv och orden på *-het* kan inte konstrueras på samma sätt som sin finska motsvarighet. Vi kan jämföra några par:

| | | | |
|-----|-----------------|------|-----------------------|
| Ex. | kiltti | -ys | kiltteys el. kilttiys |
| | snäll | -het | snällhet |
| | aineellinen-uus | | aineellisuus |

| | |
|-----------------|--|
| materiell | *materiellhet |
| materialistisk | materialism |
| aiheellinen-uus | aiheellisuus |
| befogad | *befogadhet |
| välgrundad | *välgrundadhet |
| motiverad | *motiveradhet |
| kutistua > | |
| kutistuva-uus | kutistuvuus |
| krympa > | |
| krympande | *krympandehet 'benägenhet att krympa' |
| nyvä > | |
| parempi-uus | paremmuus |
| bra > | |
| bättre | *bättrehet |

Också den här typens substantiveringar ingår frekvent i nominalfraser på finska:

- Ex. epäilyksen aiheellisuus
- det befogade i misstanken
"att misstanken var befogad"
- hänen olemuksensa salaperäisyys
- det hemlighetsfulla i hans väsen
- eller som motsvarighet i löpande text
"hans hemlighetsfulla väsen"

2.3. Personbetecknande substantiv

Samma problematik som ovan återkommer vad gäller de personbetecknande substantiven. I modern finska är den vanligaste personbetecknande avledningsändelsen *-ja/-jä*. Den motsvaras i svenska av *-are*, men medan de svenska bildningarna i regel är beteckningar för personer som vanemässigt utför en viss verksamhet (Thorell 1981, s. 103),

kan den finska bildningen användas också om personer som tillfälligt utför något.

| | | |
|-----|---------|---|
| Ex. | ajaa | ajaja |
| | föra | förare |
| | köra | *körare |
| | juosta | juoksija |
| | löpa | löpare |
| | springa | springare (dock inte personbetecknande) |
| | avata | avaaja |
| | öppna | öppnare 'konservöppnare' *mötets öppnare |
| | alkaa | alkaja |
| | börja | *börjare |

För ett uppslagsord som *avaaja* ges liksom för *lento* de specialiserade betydelserna i själva artikeln, medan den allmänna betydelsen 'en person som öppnar något' kan hänföras till en modell.

2.4. Kausativa verb

Den sista problemtypen jag skall ta upp är kausativa verb. Det finns flera typer, som alla är relativt frekventa i finskan. På svenska motsvaras verben i princip av verbfraser som *låta göra ngt, ha ngn att göra ngt, få ngn/ngt att göra ngt, ha ngt att göras*.

| | | |
|-----|------------|-----------------------------|
| Ex. | kirjoittaa | kirjoituttaa |
| | skriva | låta skriva, ha att skrivas |
| | hoitaa | hoidattaa |
| | sköta | |
| | sköta om | låta sköta om, få omskött |
| | leikata | leikkauttaa, leikkuuttaa |
| | klippa | låta klippa, ha att klippas |

I löpande tal eller skrift väljer man dock ofta andra sätt att uttrycka saken på svenska.

| | |
|-----|--|
| Ex. | Hän leikkautti tukkansa |
| | Hon lät klippa håret Hon klippte sig |
| | Hän kävi kampaajalla leikkauttamassa tukkansa |
| | *Hon var till frissan och lät klippa håret Hon var till frissan och klippte sig |
| | Hän lähti sairaalaan hoidattamaan haavan |
| | ?Hon åkte till sjukhuset för att låta sköta om såret Hon åkte till sjukhuset för att få såret omskött |

Modellen för den här typen kommer att få en något annan funktion än för de tidigare beskrivna typerna. Gemensamt är här de mera teoretiska motsvarigheterna, medan sättet att uttrycka sig i löpande skrift eller tal i högre grad varierar från verb till verb och därför måste tas upp under själva uppslagsordet. Fördelen med en modell för typen är att alla de verb som inte har andra motsvarigheter än de bildningsmässiga motsvarigheterna *låta göra, ha att göra* o.s.v. kan behandlas med enbart en hänvisning. Modellen skall därtill ge exempel på vissa mönster som trots allt finns:

| | | |
|-----|--------|-----------|
| Ex. | paisua | paisuttaa |
|-----|--------|-----------|

Det finska ordet *paisua* har betydelser som 'svälla, svullna; växa [i omfång]; stiga (om vatten); expandera, öka' och i alla de här betydelserna kan det ha en kausativ motsvarighet *paisuttaa*. En del exempel på användningen är verbtypiska och måste ges i artikeln *paisuttaa* men ett mönster som återkommer för många verb och som kommer att behandlas i modellen är att svenskan ofta s.a.s. vänder på steken. I stället för att säga *regnet fick älven att stiga* eller *regnet gjorde att älven steg* kan man säga *älven steg på grund av regnet*. En motsvarande konstruktion förekommer på finska men den anses ofinsk.

3. Modellartiklarna

Översättningsmodellerna finns tillsvidare inte utarbetade, men riktlinjerna är i stort sett klara. Varje modell kommer att bestå av en kortfattad explicit förklaring av problematiken och en exempelsamling där vi ger prov på vilka syntaktiska motsvarigheter typen kan motsvaras av på svenska. Exempelen kommer att hämtas från flera olika ord för att

modellen skall bli allsidig.

Ursprungligen hade vi tänkt oss att modellerna skulle placeras i anslutning till förordet, men senare har vi kommit fram till att de skall placeras som informationsrutor inne i själva ordboken så att användaren stöter på dem i sitt bläddrande. I samband med innehållsförteckningen kommer det naturligtvis att finnas en förteckning över dem. Modellerna måste göras klara och överskådliga så att användaren känner sig lockad att läsa dem.

Från själva ordartiklarna kan vi hänvisa till typmodellen för ett problematiskt ord. Vissa ord, där ekvivalenter helt saknas på svenska, kan förbigås med en hänvisning till respektive modell. I andra fall används hänvisningen som komplement till behandlingen i själva ordartiklen.

På det här viset vill vi försöka peka på en problematik som borde vara självklar för alla som sysslar med översättning men som tyvärr inte alltid är det — att det inte råder något ett till ett-förhållande mellan källspråkets och målspråkets uttrycksmedel.

Litteratur

Ikola, Osmo. 1986. *Nykysuomen käsikirja*. Weilin+Göös. Espoo.

Reuter, Mikael. 1993. *Översättning och språkriktighet*. Svensk Språktjänst AB. Tryckericentralen. Helsingfors.

Suomen kielen Perussanakirja L-R. 1992. Kotimaisten kielten tutkimuskeskuksen julkaisu 55. Valtion painatuskeskus. Helsinki.

Thorell, Olof. 1973. *Svensk grammatik*. Esseltestudium/ Norstedt & söner. Stockholm.

Lemmasektion til Nordisk Leksikografisk Ordbog

Hans Kristian Mikkelsen

1. Indledning

Nærværende artikel er én af tre i denne bog, som behandler teoretiske og praktiske leksikografiske problemer i forbindelse med udarbejdelsen af en ordbog, som bærer arbejdstitlen "Nordisk leksikografisk ordbog" (NLO). De to andre er Ruth Vatvedt Fjelds og Bo Svenséns, som uddyber henholdsvis ækvivalentproblematik og fagsystematik. En foreløbig version af NLO blev rundsendt til medlemmerne af Nordisk Forening for Leksikografi (NFL) forud for konferencen. Det udsendte materiale var udelukkende tænkt som en status over projektets hidtidige arbejde samt som diskussionsmateriale og foregiver på ingen måde at være en færdig ordbog.

2. NLOs historie

Straks fra starten af sit virke satte bestyrelsen for NFL en ordbog over leksikografisk terminologi på sin ønskeseddel, og på formanden Dag Gundersens initiativ blev der i løbet af efteråret 1991 sammensat en planlægningsgruppe, som fik til opgave at udarbejde et koncept til en nordisk leksikografisk ordbog. Jeg vil ikke her komme nærmere ind på dette forløb, som man iøvrigt kan læse om andetsteds¹, men vil nøjes med at konstatere, at der på et møde i Oslo i februar 1992 blev udviklet et sådant koncept. Derpå blev gruppen udvidet, således at flg nordiske sprog var repræsenteret: dansk, finsk, islandsk, norsk og svensk. Siden har finnerne trukket sig ud, men vi håber da, at de vender tilbage. Det er også grunden til, at finsk er bevaret i den rundsendte version.

Redaktionsgruppens arbejdsform er, tror jeg, et typisk eksempel på nordisk samarbejde. Her lægges der vægt på demokrati med alle de fordele og ulemper, det kan indebære. Der er ingen indianere: alle er høvdinge, og det er ikke ubetinget et gode, når det gælder udarbejdelsen af en ordbog. Foreløbig fungerer det, men der er situationer, hvor det vil være ikke blot lettere, men måske også bedre mhp det endelige produkt, at have en redaktør, der tager beslutninger i tvivlstilfælde og på den måde sikrer, at ordbogen får et ensartet præg. Jeg skal senere vende tilbage til sådanne eksempler.

Retningslinjerne for arbejdet udstikkes på møder, hvor alle deltager. ("Alle" har hidtil betydet: to fra Danmark, en fra Finland, en fra Island, to fra Norge og en fra Sverige.) Det praktiske arbejde med ordbogen foregår decentralt, dvs i de enkelte lande. Redigeringsarbejdet foregår ved en Macintosh-computer, idet ordbogen er lagret i databaseprogrammet FileMaker Pro. Undertegnede sørger for at administrere databasen, der kræver forskellige koder, for at man kan komme til den. Den enkelte medarbejder har

kun adgang til oplysningerne om sit eget sprog. Større ændringer såsom tilføjelse eller sletning af kort kræver en kode, som kun jeg kender. På denne måde sikres det, at der ikke går kludder i systemet med udsendelse og returnering af disketter fra, respektive til, Århus.

Leksikografiens fagsprog er ikke tidligere selv blevet udsat for leksikografisk behandling². Undersøger man diverse lingvistiske ordbøger, opdager man, at leksikografien er yderst sparsomt repræsenteret, hvilket vel skyldes, dels at metaleksikografien først i de seneste par årtier har markeret sig, dels at beskæftigelse med ordbøger traditionelt ikke har været betragtet som videnskabeligt *comme il faut*. Man finder heller ingen videre hjælp i de lingvistiske fagordbøger, når det gælder ideer til strukturering, idet det generelt gælder for lingvistiske ordbøger, at de er monolingvale. En fælles nordisk leksikografisk ordbog må nødvendigvis være polylingval. Til gengæld findes der ganske mange andre eksempler fra andre fagområder på sådanne ordbøger.

Vi har i planlægningsgruppen fra begyndelsen opereret med syv funktioner eller formål, som man kunne tænke sig varetaget af en nordisk leksikografisk ordbog³. De drejer sig om to monolingvale: reception og produktion på modersmål, fire bilingvale: reception og oversættelse fra fremmedsprog, produktion og oversættelse til fremmedsprog, samt én encyklopædisk funktion: indsigt i fagområdet. Det bedste ville selvfølgelig være, hvis man kunne tilgodese alle disse funktioner optimalt i én og samme ordbog. Men det kan ikke lade sig gøre, eftersom de stiller vidt forskellige krav til både selektion og præsentation. Et af forslagene til koncept for NLO gik således ud på at sætte reception af tekster på engelsk, fransk og tysk øverst på en prioriteringsliste, og den naturlige konsekvens heraf var at gøre ikke-nordiske tekster til empirisk basis, altså til det tekstgrundlag, hvorfra lemmaerne skulle vælges. Og for præsentationen ville det betyde, at den optimale løsning lå i en alfabetisk makrostruktur med engelske, franske og tyske lemmaer. Ville man derimod prioritere den faglige indsigt, altså terminologien, højest, ville den optimale løsning være at begynde med udarbejdelse af et begrebssystem. En sådan løsning ville medføre en meget stram lemmaselektion og muligvis præsentation i en systematisk makrostruktur.

Resultatet af overvejelserne i planlægningsgruppen blev, at nordiske leksikografer som ordbogens målgruppe antoges at have brug for et hjælpemiddel til produktion af faglige tekster på deres modersmål og til reception af engelske, fransk og tyske leksikografiske tekster. Siden er hensynet til de udenlandske sprog blevet prioriteret endnu lavere, således at vi idag lader tre funktioner være styrende for vores arbejde. Disse tre er indbyrdes prioriteret:

produktion på modersmål
reception på modersmål
fagforståelse

De øvrige fire funktioner spiller højst en sekundær rolle, og man kan dårligt tale om en indbyrdes prioritering af disse, hvorfor de her anføres i alfabetisk orden:

oversættelse fra fremmedsprog
 oversættelse til fremmedsprog
 produktion på fremmedsprog
 reception på fremmedsprog

Hvilke følger fik denne prioritering så for selektion og præsentation? Det optimale udgangspunkt ville uden tvivl være, hvis man kunne sammensætte en empirisk basis af leksikografiske tekster, der var affattet på de nordiske sprog. Udvalget af egentlig leksikografiske tekster på disse sprog er imidlertid så sparsomt, at det i bedste fald ville være mangelfuldt og utilstrækkeligt. Derfor valgte vi den næstbedste løsning, nemlig at gøre ikke-nordiske tekster til kernen i den empiriske basis. Denne kerne blev suppleret med materiale fra eksisterende nyere tekster på nordiske sprog: dels Bo Svenséns håndbog fra 1987, dels rapporten fra NFLs første møde i Oslo for to år siden. Dertil kommer, at forordene i mange nordiske ordbøger er kigget igennem mhp excerpering. I praksis blev det dog ikke primært teksterne selv, men den essens af dem, som findes repræsenteret i registre, der kom til at udgøre kildematerialet. Således valgte vi at gøre registeret i den store håndbog i leksikografi, Hausmann et al. 1989-1991, til hovedkilden.

Om følgerne for præsentationen vil jeg koncentrere mig om det, som har at gøre med lemmaselektionen, nemlig makrostrukturen. Omvejen ad en empirisk basis, der fortrinsvis bestod af engelske, franske og tyske tekster/ord, betød, at der først måtte foretages en oversættelse til nordisk af de selekterede udenlandske ord, før man kunne få et nordisk lemma frem. Som det fremgår af den rundsendte version af NLO blev bokmål fundet bedst egnet til at kunne fungere som lemmasprog. Der var nu i hvert fald to måder at påbegynde arbejdet. Efter at have udvalgt de engelske, franske og tyske ord kunne man *enten* oversætte dem til bokmål, for siden at tilskrive ækvivalenter fra de øvrige sprog - *eller* oversætte dem til alle de nordiske sprog, for siden at gøre bokmålsækvivalenten til lemma. Vi valgte den sidstnævnte løsning. Fordelene ved den var, udover at arbejdet kunne komme igang straks i alle lande, at ansvaret for lemmaernes gyldighed og relevans blev fordelt på alles skuldre. Vi er nemlig ofte i den situation at skulle være normdannende på området, især fordi vi ofte støder på ord, som ikke har nogen entydig og veletableret pendant på vores sprog. Sagen er jo den, at med den arbejdsmodel, som jeg netop har skitseret, vil den del af ækvivalensproblematikken, som handler om at harmonisere de nordiske udtryk, også blive bestemmende for, hvordan lemmaet kommer til at se ud⁴.

I den første del af arbejdet har ordbogen således været uden noget egentligt lemmasprog, eftersom ikke blot engelske, franske og tyske, men også danske, norske og svenske ord er blevet selekteret. Først fornylig er bokmål blevet gjort til lemmasproget, hvilket den foreløbige version også bærer præg af. Bortset fra at der stadig kan fyldes nye lemmaer på - og fjernes nogle, som viser sig ikke at leve op til de kriterier, som jeg vil komme ind på nedenfor - vil arbejdet i den nærmeste fremtid blive koncentreret om at rydde op i bokmålslemmaerne. For naturligvis går det ikke at have flere forskellige ord som indgang til samme artikel. Grunden til at sådanne uacceptable lemmaer findes på nuværende tidspunkt er altså ikke sjusk fra vores side, men en følge af den beskrevne arbejdsgang

med engelsk, fransk og tysk som typisk indfaldsvinkel.

3. Lemmaselektion til NLO

Dette kapitel vil dreje sig for det første om det, der er fælles for de udvalgte ord, og for det andet om det, der adskiller dem og motiverer en mikrostrukturel forskelsbehandling. Strengt taget er det kun det første punkt, der er relevant for lemmaselektionen, mens det andet mere præcist indebærer en klassifikation af det udvalgte materiale. Imidlertid er det de samme kriterier, der bruges som udgangspunkt for selektion og klassifikation. En fælles behandling motiveres yderligere af den omstændighed, at der måske er behov for en ny selektion, som kan udføres sammen med klassifikationen. I det hele taget er det, som det allerede er fremgået, et spørgsmål, om betegnelsen "lemmaselektion" er dækkende for det hidtil udførte arbejde. Jeg vil dog holde fast i den i det følgende.

3.1. Selektion: det der forener

Det er ret begrænset, hvad der er skrevet specielt om lemmaselektion til fagordbøger, ja såmænd også til ordbøger i det hele taget. Der er nok traditionelt snarere tale om en praksis, et håndværk end et teoretisk afklaret fænomen. Jeg ser her bort fra de kommentarer, man finder i ordbogsanmeldelser, som ofte fokuserer på selektionsspørgsmål, men på en atomistisk måde, hvor en ordbog typisk kritiseres for ikke at have optaget nogle bestemte ord.

Ikke desto mindre vil jeg indlede med at trække nogle teoretiske linjer op, så meget mere som man med en vis ret kan sige, at vi stadig - eller først nu - befinder os i denne fase. Under hele det hidtidige arbejde har vi taget udgangspunkt i fem principper for lemmaselektion, formuleret af Henning Bergenholtz (under udg.b). De fem principper er

- (1) kun termer
- (2) termer og andre udtryk, som bruges anderledes end i almensproget
- (3) termer og andre udtryk, som er særligt vigtige eller relevante for et givet fagsprog
- (4) udtryk, som er særligt regelmæssige i et givet fagsprogs tekster ("skjult frekvens")
- (5) udtryk, som optræder med en vis frekvens i et givet fagsprogs tekster

Femdelingen dækker efter min opfattelse over nogle interne ligheder og forskelle, som ville kunne motivere en lidt anden ordning. For det første skjuler der sig mellem 1-3 på den ene side og 4-5 på den anden en modsætning mellem en deduktiv og en induktiv tilgang. Denne forskel afspejler to forskellige fagsprogsforståelser: den deduktive tager udgangspunkt i faget, den induktive i de tekster, der produceres og reciperes inden for et fag. Den deduktive er vidensbaseret - og den induktive tekstbaseret. I sin rene form kan det induktive princip ikke sikre, at der overhovedet vil blive optaget fagord og fagtermer i en ordbog baseret på fagtekster - ligesom det deduktive princip heller ikke kan

garantere, at de udvalgte fagord også vil være dem, man finder i fagteksterne. For det andet er der en skillelinje mellem 1-2 og 3. 1-2 holder sig inden for det, der kendetegner et enkelt fagområde, mens 3 ikke udelukker, at der kommer stof fra andre fagområder med. Endelig er der, for det tredje, en forskel mellem "strammere" og "slappere" i forholdet mellem 1 og 2-3, samt mellem 4 og 5. I den induktive tilgang er strammerne dem, der går ind for at anvende objektive frekvenskriterier fremfor at slå sig til tåls med den intuitive, subjektive sprogkompetens⁶. I den deduktive tilgang er strammerne dem, der tænker i rent terminologiske baner. For dem står begrebet i centrum, og kun de ord, der på indholdssiden objektivt modsvares af et fagligt begreb, kan selekteres. For de deduktive slappere spiller det en mindre rolle, om der strengt taget er tale om en term. Ved denne mere subjektive tilgang er synonymer, konnotationer, varianter mm ikke tabu - i modsætning til hvad der er gældende for den ortodokse terminologiske forståelse⁷. Altså synes de fem principper at fordele sig på tre dimensioner, som det fremgår af nedenstående illustration:

| | induktiv, tekstbaseret | deduktiv, vidensbaseret | |
|-----------|---------------------------|-----------------------------|-----------------------------|
| | | faggrænse- respekterende | faggrænse- overskridende |
| objektiv | 5 | 1 | |
| subjektiv | 4 | 2 | 3 |

(Tallene refererer til de fem ovenfor anførte principper. Det tomme felt står for et kriterium, som man også meget vel kunne tænke sig anvendt, nemlig når det gælder en strengt terminologisk ordbog, der dækker mere end ét fag.)

Hvad har vi så gjort - i teori og i praksis? Under ideelle omstændigheder vil valget af kriterium bero på formålet med ordbogen. Som allerede nævnt er det hensigten, at nordiske leksikografer som ordbogens målgruppe skal have et redskab, der hjælper dem ved produktion og reception af tekster på modersmål samt giver en indsigt i faget leksikografi. Hvis det kun havde været vores hensigt at give oplysninger om faget leksikografi, ville kriterium 1 have været et oplagt valg. Diskussionen stod mellem 2 og 3⁸. Argumenterne for at bruge 3 var primært af pædagogisk karakter. Ved undervisning i leksikografi behandles også emner af mere almen lingvistisk karakter, som de studerende også har brug for at kunne søge oplysninger om. Vi afprøvede dette kriterium på grundlag af en liste over lemmakandidater fra Crystal 1991, og anvendt på et tilfældigt udsnit af bogstavet "i" førte det til, at følgende ord måtte anses for relevante for NLO:

infix
informant
initial
International Phonetic Association (IPA)
intonation
intransitivity
irregular

Konklusionen var, at det fagoverskridende selektionsprincip reelt ville gøre ordbogen til en lingvistisk snarere end en leksikografisk ordbog. Det endte med, at vi besluttede os for kriterium 2, der kendetegnes ved at holde sig inden for leksikografiens rammer i snævrere forstand, men uden at begrænse selektionen til egentlige termer⁹. Til gengæld for at gøre selektionen forholdsvist bred, sådan som den subjektive slapperholdning er udtryk for, begrænsede vi også selektionens dybde. Ved en "dyb selektion" forstår jeg en selektion, som bevæger sig relativt langt ned i det fagsystematiske hierarki og ud i de faglige forgreninger. I praksis betyder det, at vi som hovedregel ikke tager enheder på mere end to ord med. En helt kontant følge af dette ledsagende princip er, at mange af Wiegands innovationer på den leksikografiske terminologis område ikke kommer med - i hvert fald ikke som lemmer. Fra den nedenstående liste, der rummer forskellige forgreninger af termen "Mikrostruktur" (M.)¹⁰ kommer således kun de understregede med:

Mikrostruktur

M. mit aufgespaltener Kernstruktur
M. mit binnenerweiterter Basisstruktur
M. mit linkerweiterter Basisstruktur
M. mit rechtserweiterter Basisstruktur
abstrakte M.
abstrakte lineare M.
annexierte M.
annexierte hierarchische M.
auflistende M.
auflistende M. mit linkerweiterter Basisstruktur
einfach subintegrierte M.
einfach integrierte M.
einfach integrierte hierarchische M.
einfach vollständig erweiterte M.
extern vollständig erweiterte M. mit linkerweiterter Basisstruktur
hierarchische M.
integrierte M.
integrierte M. mit binnenerweiterter Basisstruktur
integrierte rechtserweiterte M.

intern erweiterte M.
konkrete M.
linkserweiterte M. mit links- und rechtserweiterter Basisstruktur
nichtintegrierte M.
partiell auflistende M.
partiell integrierte M.
partitive M.
präzedentive M.
rechtserweiterte M.
rechtserweiterte integrierte M.
rechtserweiterte M. mit binnenerweiterter Basisstruktur
rechtserweiterte nichtintegrierte M.
rudimentäre M.
semiintegrierte M.
überdachende M.
überdachende nichtrudimentäre M.
überdachende partiell rudimentäre M.
überdachende vollständig rudimentäre M.
zusammengesetzte M.
zweifach partiell integrierte M.
zweifach subintegrierte M.
Mikrostrukturenprogramm

Til sammenligning ser NLOs bestand af bokmålslemmaer, hvori "mikrostruktur" (m.) indgår, således ud:

abstrakt m.
anneksert m.
enkel m.
hierarkisk m.
integrert m.
konkret m.
lineær m.
mikrostruktur
mikrostrukturprogram
obligatorisk m.
oppregnende m.
overdekkende m.
overdekket m.
partitiv m.
presedentiv m.

rudimentær m.
sammensatt m.
utvidet m.

I praksis har det hidtil vist sig at være vanskelig at gøre selektionen konsistent i forhold til det valgte hovedkriterium (2). De problemer, som den enkelte kan have med at være konsekvent i sin egen måde at selektere, forstærkes af de problemer, som opstår, når man beslutter sig til at gøre arbejdet på ægte demokratisk vis¹¹. Desuden begyndte vi som allerede omtalt ikke med at udarbejde en systematisk fagoversigt, som ellers ville have været en hjælp i disse tvivlstilfælde, som nu måtte løses per intuition. Reelt betyder det, at det materiale, som foreligger på tryk i den foreløbige version, er en blanding af kriterierne 2 og 3. I bakspejlet kan man således konstatere, at det måske havde været bedre at begynde med at udarbejde en faglig systematik og dermed tilvejebringe et mindre subjektivt beslutningsgrundlag.

3.2. Klassifikation eller anden selektion: det der skiller

De 1401 indgange, som NLO har nu, udgør således en noget blandet forsamling. Jeg har nævnt to hovedårsager hertil: for det første at vi har måttet ty til ikke-nordisk kilde-materiale og for det andet at det har været vanskeligt at overholde det valgte selektions-kriterium.

Derfor kan man godt kalde den klassifikation af materialet, som er næste punkt i projektet, for endnu en selektion. Klassifikationens opgave er at finde frem til de centrale ord, som skal varetage ordbogens faglige, encyklopædiske intention¹². Hensigten er at føje en terminologisk dimension til den leksikografiske. Dog vil det leksikografiske aspekt fortsat have førsteprioritet, hvilket både fremgår af valget af kriterium 2 som det vigtigste og beslutningen om at ordbogen skal have en alfabetisk og ikke systematisk makrostruktur. Det betyder, at man finder synonyme, stilistiske varianter mm på deres alfabetiske plads i ordbogens makrostruktur. Fra disse i terminologisk forstand perifere ord vil der gives henvisninger til centrale ord, "klasse 1-ord", således at man dér kan få en forklaring, hvis man søger en sådan. Meningen er, at der skal henvises opad i systemet, altså fra et perifert til et centralt leksikografisk ord, men ikke omvendt. Klasse 2 omfatter ikke nødvendigvis kun stilistisk markerede leksikografiske fagord. Man kan udmærket tænke sig, at der fra toleddede ikke-idiomatiske lemmaer henvises til enleddede for en samlet forklaring, hvilket vil placere de toleddede i klasse 2. Fx:

| | | |
|---------------------------|----|--------------------|
| aktiv oversettelsesordbok | -> | aktiv ordbok |
| artikkelinndeling | -> | mikrostruktur |
| enspråklig ordbok | -> | monolingval ordbok |
| stikkord | -> | lemma |
| artikkelpresentasjon | -> | presentasjon |
| ordboksartikkel | -> | artikkel |

| | | |
|------------------------|----|---------------|
| abstrakt mikrostruktur | -> | mikrostruktur |
| enkel mikrostruktur | -> | mikrostruktur |

Spørgsmålet er så, hvad der skal gøres med restgruppen, ordene fra klasse 3, som jo er kommet med ad bagvejen. Efter min fortolkning af selektionskriterium 2 er der ca 250 af de 1401 opslag, der tilhører klasse 3 - og som kun burde være optaget, hvis kriterium 3 havde været valgt. Lad mig her nøjes med at anføre sådanne eksempler med bogstavet "e":

EFL-ordbok
 egenskapsnavn
 egentlig betydning
 ekkoord
 ellipse
 etnosentrisitet (i ordbøger)
 etterledd
 etymologi
 etymon, opphavsord
 eufemisme

Der er i princippet to muligheder. Hvis man ikke ændrer på de oprindeligt vedtagne optagelseskriterier, må klasse 3-ordene slettes. Den anden løsning, som måske er den mest realistiske, hvis man forudsætter, at der ikke ændres på den demokratiske arbejdsproces, er, at man alligevel beslutter sig for at selekttere efter kriterium 3. Men det bør samtidig medføre, at man tager en ny selektionsomgang fra den empiriske basis, således at alle bevidst kan vælge efter kriterium 3.

4. Afslutning

Den selektion, der er foretaget fra den primært ikke-nordiske empiriske basis, kan af flere grunde ikke kaldes en ægte lemmaselektion. For det første er de nuværende bokmålslemmer ikke uddraget direkte fra en empirisk basis, men næsten udelukkende fremkommet ved oversættelse fra de forlæg på engelsk, fransk og tysk, som udgør ord-bogens egentligt selekterede materiale. Af samme grund kan man næppe heller sige, at vi har været igennem en egentlig ækvivalenttilskrivningsfase. For det andet svarer NLOs indhold på nuværende tidspunkt ikke til de vedtagne selektionskriterier. Min konklusion er, at lemmaselektionen og ækvivalentsøgningen egentlig først skal igang nu, selvom der naturligvis i de to "pseudo-" eller "kvasi-" faser allerede er gjort en del af det nødvendige arbejde¹³.

Noter

- 1 Se Bergenholtz under udg.a.
- 2 Link & Schaefer 1989.
- 3 Bergenholtz under udg.a.
- 4 For en uddybning af ækvivalentproblematikken henvises til Ruth Vatvedt Fjelds artikel i denne bog.
- 5 "Almensprog" skal i denne forbindelse forstås som omfattende alt det, der ikke indgår i det leksikografiske fagsprog, altså også andre fagsprog.
- 6 Bergenholtz (under udg.b) bemærker rigtigt, at det subjektive element, der er knyttet til sprogkompetensen, reelt blot flyttes fra lemmaselektionen til selektionen af tekster til den empiriske basis: "Zugegebenermaßen stellt eine solche Grundbestimmung von Gemeinsprache bzw. von einzelnen Fachsprachen nur eine Verschiebung des Problems dar, indem die Prinzipien für die Textauswahl zum entscheidenden Kriterium gemacht wird."
- 7 Den russiske fagleksikograf Aleksandr Gerd (1986:46) giver udtryk for slapperholdningen, når han siger, at det ikke kan være en termordbogs opgave at diskutere, hvad en term er. "Termordbogens opgave er at samle og definere såvel de egentlige termer inden for et vidensområde som alle andre sproglige udtryksformer, der på indholdsplanet hælder mod termene." (min oversættelse, HKM). Omvendt kan V.M. Pererva (1976) siges at repræsentere strammerne ved at sætte netop termafgrænsning øverst på listen over lemmaselektionsproblemer.
- 8 At de induktive kriterier slet ikke har været på tale, skyldes utvivlsomt, at disse kun egner sig til at være selektionskriterium, når det gælder ordbøger, der involverer almensproget, herunder fagordbøger for handelskorrespondance mm. Jf iøvrigt Bergenholtz under udg.b.
- 9 Blandt disse er også sådanne, som er homonymer til termer fra andre fagsprog. "Mikrostruktur" og "makrostruktur" er eksempler herpå.
- 10 Termerne er hentet fra indekset til Hausmann et al. 1989-1991, der som nævnt udgør kernen i NLOs empiriske basis. Dette indeks er både et emne- og et ordregister, og den liste, jeg har opstillet her, kan betragtes som et forsøg på at sortere emneordene fra og dermed fremstille en ren termliste.
- 11 Arbejdsgangen er beskrevet i Bergenholtz under udg.a.
- 12 Jf Bo Svenséns artikel i denne bog.
- 13 Jf iøvrigt Svenséns bidrag (under udg.), som konkluderer, at arbejdet højst kan siges at have overstået term- og begrebsinventeringsfaserne, og at det først er nu, at den egentlige systematiske selektionsfase kan begynde.

Litteratur

Bergenholtz, Henning under udg.a: "Plan eines nordischen Wörterbuches der Lexikographie. Ein Werkstattbericht" i *Symposium on Lexicography VI*, (i *Lexicographica. Series Maior*), Tübingen.

- Bergenholtz, Henning under udg.b: "Fachsprache und Gemeinsprache: Lemmaseleksjon im Fachwörterbuch" i Bergenholtz, Henning og Burkhard Schaefer (red.): *Das Fachwörterbuch. Bausteine zur Theorie und Praxis der Fachlexikographie*, Narr: Tübingen.
- Bergenholtz, Henning, Ilse Cantell, Ruth Vatvedt Fjeld, Dag Gundersen, Sigurður Jónsson, Hans Kristian Mikkelsen, Bo Svensén 1993: *Nordisk Leksikografisk Ordbog (foreløbig version, april 1993)*, Nordisk Forening for Leksikografi: Oslo.
- Crystal, David 1991: *A Dictionary of Linguistics and Phonetics*, 3rd ed., Blackwell: Oxford.
- Fjeld, Ruth Vatvedt under udg.: "Ekvivalensproblematikk i Nordisk leksikografisk ordbok", i *denne bog*.
- Gerd, Aleksandr Sergeevic 1986: *Osnovy naucno-techniceskoj leksikografii*, Izdatel'stvo Leningradskogo Universiteta: Leningrad.
- Hausmann, Franz Josef, Oskar Reichmann, Herbert Ernst Wiegand og Ladislav Zgusta (red.) 1989-1991: *Wörterbücher - Dictionaries - Dictionnaires. Ein internationales Handbuch der Lexikographie - An International Encyclopedia of Lexicography - Encyclopédie internationale de lexicographie* 1 (1989), 2 (1990), 3 (1991). De Gruyter: Berlin, New York, 1989-1991.
- Link, Elisabeth og Burkhard Schaefer 1989: "Fachsprache der Lexikographie", i Hausmann, Franz Josef, Oskar Reichmann, Herbert Ernst Wiegand og Ladislav Zgusta (red.): *Wörterbücher - Dictionaries - Dictionnaires. Ein internationales Handbuch der Lexikographie - An International Encyclopedia of Lexicography - Encyclopédie internationale de lexicographie* 1. De Gruyter: Berlin, New York, side 312-322.
- Pererva, V.M. 1976: "O principach i problemach otbora terminov i sostavlenija slovnika terminologiceskich slovarej", i Barchudarov, S.G., V.P. Petuskov og F.P. Sorokoletov (red.): *Problematika opredelenij terminov v slovarjach raznykh tipov*, Nauka: Leningrad, side 190-204.
- Svensén, Bo 1987: *Handbok i lexicografi. Principer och metoder i ordboksarbetet*, Esselte Studium: Stockholm.
- Svensén, Bo under udg.: "Lexikografins terminologi: systematiske aspekter", i *denne bog*.

Applicering af alordbogsprincippet på konkrete ordbogsartikler

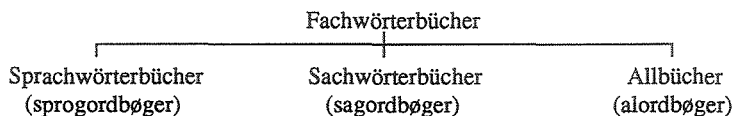
Sandro Nielsen

1. Indledning

En tosproget fagordbog har traditionelt været defineret og udarbejdet som en liste af ord, ordnet efter et bestemt princip, normalt alfabetisk, med dertil hørende ækvivalenter. På basis af de senere års metaleksikografiske forskningsarbejder kan man konkludere, at det ofte ikke er tilstrækkeligt med sådanne tosprogede ordlister. For at være så anvendelig som mulig, skal en fagordbog være udarbejdet for en nærmere defineret målgruppe og være udarbejdet til at opfylde et nærmere angivet formål. Dette vil ofte medføre, at en tosproget fagordbog bl.a. skal have en veludviklet mikrostruktur. Alt for ofte ser man, at de enkelte artiklers informationer ikke er optimale for eller tilstrækkelige til at opfylde målgruppens leksikografiske behov, fx fordi fagordbogens forfatter(e) har defineret ordbogens målgruppe så bredt, at den omfatter snart sagt alle og enhver. I det følgende vil der blive givet eksempler på denne utilstrækkelighed, og på hvilke oplysninger der kan medtages for at dække målgruppens leksikografiske behov. Dette omfatter både en ordbogstypologisk undersøgelse og en diskussion af både artikelinterne og artikeleksterne oplysninger.

2. Typologisk ramme

Der er i tidens løb gjort mange forsøg på at typologisere ordbøger, og ordbogstypologier er også stærkt afhængige af de kriterier man opstiller (jf. Zgusta 1971, Malkiel 1975, Kromann/Riiber/Rosbach 1984, Hausmann 1989 og Nielsen 1990). Det er altså muligt at fremlægge mange forskellige bud på hvordan ordbøger kan typologiseres, men her vil jeg begrænse mig til følgende overskuelige ordbogstypologi, med særlig relevans for fagordbøger:



Fagordbogstypologisk oversigt efter Wiegand 1988: 762.

Ved en **sprogordbog** forstås en ordbog, hvis genuine formål er at give oplysninger om et sprogs lingvistiske egenskaber og konventioner, altså sproglige oplysninger. Ved en **sagordbog**, også kaldet en **encyklopædi**, forstås en ordbog, hvis genuine formål er at give oplysninger om den ekstralingvistiske verden. Altså encyklopædiske oplysninger. I denne sammenhæng skal en **alordbog** opfattes som en hybrid ordbogstype. Dens genuine formål er at give både sproglige og encyklopædiske oplysninger (jf. Wiegand 1988: 762). I en tosproget alordbog finder man således sproglige og encyklopædiske oplysninger om både lemmata og ækvivalenter.

3. Leksikografisk udgangspunkt

Nu har vi set, at en ordbog bør opbygges efter et nærmere angivet ordbogskoncept og vi har set nærmere på en fagordbogstypologi. I det følgende skal vi se på det førstnævnte ordbogskoncept i forbindelse med den nævnte ordbogstypologi, dvs. forholdet mellem et komplekst, men sammenhængende ordbogskoncept og de typer oplysninger, som er relevante i en tosproget fagordbog. Dette er dog et meget vidt begreb, da tosprogede fagordbøger kan have mange forskellige formål eller funktioner og være udarbejdet for mange forskellige typer brugere. For at begrænse diskussionen, vil jeg koncentrere mig om sprogparret engelsk og dansk og disciplinen juridisk sprog. Jeg vil tage udgangspunkt i nogle konkrete artikler fra en engelsk-dansk juridisk ordbog og se på hvilke typer oplysninger artiklerne indeholder. Ordbogen er Feodor Niensens *Anglo/amerikansk-dansk retsordbog*, og jeg vil tage udgangspunkt i følgende artikler hentet fra bogstavet C:

contract kontrakt, overenskomst, aftale.
contract in restraint of trade konkurrenceklausul, konkurrencebegrænsende aftale.
contract note slutseddel.
contract of adhesion adhæsionskontrakt.
contract to sell aftale om senere at overdrage ejendomsretten til et formuegode.
contractor værksmester.
contravene (v) overtræde.
contribution bidrag.
contumacy, actual nægtelse af at følge en kendelse eller beslutning truffet under en retssag.
contumacy, presumed undladelse af, trods lovlig forkyndelse, at give møde i retten.

Generelt kan man dele disse artikler i to hovedgrupper. Den første udgøres af de artikler, der består af et lemma og én ækvivalent uden yderligere oplysninger. Den anden ho-

vedgruppe udgøres af de artikler, der består af et lemma og 2 eller flere ækvivalenter uden yderligere oplysninger. Den eneste undtagelse til denne generelle opdeling er artiklen **contravene**, som foruden lemma og ækvivalent indeholder en sproglig oplysning i form af en ordklasseangivelse. Hvis man således prøver at placere FEO inden for de ovenfor skitserede typologiske rammer, må man sige, at FEO er en sprogordbog, da dens genuine formål er at give sproglige oplysninger i form af målsproglige ækvivalenter.

Hvorvidt disse oplysninger er tilstrækkelige i en engelsk-dansk juridisk ordbog afhænger af to ting. For det første er ordbogens målgruppe af væsentlig betydning, da en ordbog kun er optimal for en brugergruppe såfremt den opfylder denne brugergruppes leksikografiske behov. For det andet er ordbogens funktion vigtig, da dens funktion er med til at afgøre hvilke oplysninger, der skal inkorporeres i ordbogen. I FEO oplyses i forordet, at ordbogens målgruppe er translatørstuderende, dvs. fagligt ukyndige brugere. Ser man på denne gruppes leksikografiske behov er det klart, at ordbogen skal opfylde i hvert fald to formål: Den skal kunne anvendes til oversættelse og til tekstreception. Translatørstuderende uddannes til at oversætte, og for at kunne gøre dette bliver de nødt til at tilegne sig viden om et fremmed fagsprog gennem læsning af lærebøger. Man kan argumentere for, at der i så fald skal udarbejdes to selvstændige ordbøger, men da FEO slår disse funktioner sammen, må man tage udgangspunkt i dette.

Hvad er så målgruppens leksikografiske behov? I denne forbindelse vil jeg tage udgangspunkt i de foreliggende brugerundersøgelser på området, hvoraf nogle af de væsentligste er foretaget af Fagleksikografgruppen ved Handelshøjskolen i Århus (jf. Duvå/Laursen/Maidahl 1992). Selvom disse undersøgelser ikke kun omfatter sprogpåret engelsk og dansk, kan man dog sammenfattende sige, at brugere, der falder inden for definitionen af målgruppen for FEO har brug for tre typer oplysninger i en engelsk-dansk juridisk ordbog. Disse kan sammenfattes således:

1. Oplysninger om engelsk jura og juridisk sprog.
2. Oplysninger om dansk jura og juridisk sprog.
3. Oplysninger om flerordsforbindelser på engelsk og dansk.

4. Oplysninger om jura

I en engelsk-dansk juridisk ordbog som FEO vil oplysningerne om engelsk jura relatere sig til lemmata, medens oplysningerne om dansk jura relaterer sig til ækvivalenterne. Dette har bl.a. betydning når man skal forsøge at give disse oplysninger i en ordbogsartikel, da denne skelnen kan anvendes direkte på sekvensen af mikrostrukturelle enheder. Ved **mikrostrukturelle enheder** forstås de selvstændige oplysningsenheder, der findes i en ordbogsartikel. Alle oplysninger, der er adresserede til lemmaet placeres direkte efter dette og før ækvivalenten, mens alle oplysninger, der er adresserede til ækvivalenten placeres direkte efter denne (jf. Grindsted 1988: 32). Såfremt dette princip følges konsekvent, vil ordbogsbrugeren aldrig være i tvivl om hvad den leksikografiske adresse for

oplysningerne er. Dette princip kan illustreres med udgangspunkt i artiklen **contravene** fra FEO. Hvis denne artikel udvides således, at den opfylder de behov, der stilles af brugeren kan den have følgende udformning:

contravene v [handle således, at en krænkelse opstår]

1. **overtræde** v [person krænker en lov, kontrakt eller et forbud]
2. **stride mod** v [ting i strid med noget andet]

Som det fremgår af denne reviderede artikel, er der én faglig forklaring adresseret til lemmaet og to faglige forklaringer adresseret til hver sin ækvivalent, og i det hele taget er informationsmængden i artiklens mikrostruktur forøget betydeligt. Når artikler indeholder mange slags oplysninger er det vigtigt, at brugeren ikke mister overblikket. Derfor bør det fremgå klart af artiklens skriftbillede, hvorledes de mikrostrukturelle enheder er arrangeret, og dette kan med fordel ske ved anvendelse af **strukturindikatorer** og **leksikografiske symboler** (jf. Wiegand 1989: 428). For overskuelighedens skyld bør man placere de faglige/encyklopædiske angivelser inden for skarpe parenteser, da der således aldrig kan være tvivl om, hvor den faglige angivelse begynder og slutter. Det er her vigtigt, at der ikke kan opstå tvivl hos brugeren om hvilken adresse en faglig forklaring har, og dette undgås ved klart at angive hvor de leksikografiske adresser er placeret i en ordbogsartikel. Hvis et lemma behandles polysemt, kan man med fordel anvende **polysemiindeks** i form af arabertal, således som det er illustreret i den omarbejdede artikel. De skarpe parenteser og den fede skrift fungerer som strukturindikatorer, mens disse samt arabertallene fungerer som sekvenseringsenheder (*ordering devices*, jf. Hausmann/Wiegand 1989: 344). Dette betyder dog ikke, at man altid behøver at have forklaringer om både lemma og ækvivalent. Såfremt et lemma og dets ækvivalent er fuldstændige ækvivalenter, er det uhensigtsmæssigt og pladskrævende at angive definitioner vedrørende begge. I FEO vil det være tilstrækkeligt at angive en definition, som er adresseret til lemmaet. Grunden til, at lemmaet forklares er, at en engelsk-dansk juridisk ordbog beskriver engelsk jura og ikke dansk jura. Et eksempel på denne fremgangsmåde kan illustreres ved hjælp af følgende omarbejdning af artiklen **contract of adhesion**:

contract of adhesion n [ensidigt udformet standardkontrakt,
der kun tilgodeser den ene parts interesser]
adhæsiionskontrakt n
→ model contract; standard form contract

Som et yderligere supplement til forståelsen af engelsk jura og engelsk juridisk sprog kan man i artiklen inkorporere krydshenvisninger, her indikeret ved en pil: →. Disse kan være med til at vise brugeren, hvorledes et juridisk ord eller begreb er placeret i det juridiske fagsystem i forhold til andre ord eller begreber. I det konkrete tilfælde er *standard form contract* overbegrebet, og der findes to former heraf, nemlig *model contracts* og *contracts of adhesion*. I et udbygget krydshenvisningssystem vil en bruger af ordbogen blive gjort opmærksom på denne sammenhæng, hvis man forestiller sig, at de to øvrige artikler i den imaginære ordbog indeholder oplysninger herom (dette konkrete tilfælde kan sammenholdes med de pågældende artikler i Basisordbog eng-da). Antallet af sådanne krydshenvisninger er naturligvis forskellig fra lemma til lemma, men man bør dog for overskuelighedens skyld søge at begrænse antallet (fx til 5). Endvidere bør der kun være krydshenvisninger til vigtige lemmata, ellers vil brugeren ikke få et optimalt udbytte. Henvises der til marginalt relevante lemmata vil brugeren hurtigt opgive at benytte disse henvisninger fordi de tilsyneladende ikke er relevante og opslagene er spild af tid.

I forbindelse med ækvivalentangivelse bør man tage sig i agt for ikke at foretage en sammenblanding af en egentlig ækvivalent (*ready translation equivalent*) og den dertil hørende uddybende forklaring. Ser man på de to artikler **contumacy** kan man se, at ækvivalentangivelsen består af en **forklarende ækvivalent** eller **episk ækvivalent**. Dette kan selvfølgelig ikke altid undgås, især ikke i forbindelse med engelsk og dansk ret, da de to retssystemer er grundlæggende forskellige. Dog kan man i begge disse artikler erstatte forklaringerne med en egentlig ækvivalent og gøre den episke ækvivalent til en uddybende forklaring. I en ordbog, der er beregnet til oversættelse bør man principielt angive en ækvivalent, som kan indsættes i en oversættelse, i hvert fald i de tilfælde, i hvilke dette er muligt. Dette kan illustreres af følgende omarbejdede artikelforslag:

actual contumacy n [nægtelse af at følge en kendelse eller beslutning truffet under en retssag]
uefterrettelighed n
→ presumed contumacy
presumed contumacy n [undladelse af, trods lovligt forkyndt tilsigelse, at give møde i retten]
udeblivelse n
→ actual contumacy

I forbindelse med denne ækvivalentangivelse kan man komme ud for, at man på dansk har to ord, der er fuldstændige synonyme semantisk set. Ofte viser det sig, at der dog er en forskel på disse synonyme, idet det ene anvendes hyppigere end det andet, evt. fordi der er sket en sproglig fornyelse. For eksempel blev det danske udtryk *ansvar uden culpa* tidligere anvendt for det, der nu betegnes *objektivt ansvar*. Såfremt det er muligt

for leksikografen bør sådanne forhold tages i betragtning. Dette betyder, at det af artiklen skal fremgå hvilket af de to eller flere synonymer er det mest anvendte. Her kan man tænke sig, at den mest frekvente ækvivalent angives i **ækvivalentpositionen**, og de øvrige mindre frekvente synonymer placeres i en selvstændig **synonymiangivelse**. I en engelsk-dansk juridisk ordbog vil angivelsen af danske synonymer hjælpe brugeren i forbindelse med både tekstreception og oversættelse fra engelsk. Dette kan illustreres ved hjælp af følgende omarbejdede artikel:

contract note n [dokument, der fastslår hovedindholdet i en afsluttet handel]

1. **slutseddel** n [almindelig betegnelse inden for erhvervsliv]

2. **børsnota** n [bevis for, at aktier er blevet købt eller solgt, men endnu ikke betalt]

= fondsnota

→ bought note; broker's contract; sold note

I den omarbejdede artikel **contract note** bør man bemærke, at synonymet kun hører til den ene af de danske ækvivalenter (delbetydning 2), og den er da angivet i den artikeldel, som kun vedrører det relevante ækvivalente danske udtryk.

5. Flerordsforbindelser

Vi skal nu se nærmere på den sidste af de tre ovennævnte oplysningsgrupper. Flerordsforbindelser er vigtige i fagsproget, især da de ikke altid er transparente syntagmatiske enheder. At flerordsforbindelser er intransparente har i denne sammenhæng en dobbeltbetydning. For det første er det vigtigt at kende både de engelske og danske faste ordforbindelser for at kunne oversætte en engelsk juridisk tekst til et idiomatisk dansk. For det andet er engelske juridiske ordforbindelser ofte i sig selv uigennemskuelige betydningsmæssigt for ordbogens målgruppe, og derfor er det særligt vigtigt at medtage disse i en engelsk-dansk juridisk ordbog. De illustrerer ikke blot sproglige forskelle, men hjælper også den danske bruger til en forståelse af engelsk jura. I mikrostrukturel sammenhæng betegnes den udgangssproglige ordforbindelse **fraseangivelse**, og den mål-sproglige oversættelse betegnes **parafraseangivelse**. Et eksempel på, at FEO kan forbedres ved en inkorporering af flerordsforbindelser i en artikel kan illustreres ved følgende omarbejdede artikel:

contribution n [ydelse, der erlægges af én person til en anden]

1. **bidrag** n <pl -> [ydelse givet til støtte for et bestemt formål eller foretagende]

2. **indskud** n <pl -> [ydelse givet til sammenslutning eller foretagende, evt. betingelse for at deltage]

▲ *shares may be subscribed against contribution of other values than cash* aktier kan tegnes mod indskud af andre værdier en kontanter; *contributions shall not be in the form of contracts for work or services* indskud kan ikke betså i forpligtelser til at udføre arbejde eller erlægge tjenesteydelser

3. **regres** n [krav på at få en udgift helt eller delvis dækket af en anden]

▲ *the tortfeasor who is sued may ask for contribution from the other tortfeasors* den sagsøgte skadevolder har regres mod de øvrige skadevoldere; *no contribution can be obtained where the person sued is liable to indemnify his co-tortfeasor* der kan ikke opnås regres såfremt den sagsøgte er pligtig til at skadesløsholde medskadevolderen → civil liability contribution; right of contribution

For at gøre artiklen så brugervenlig som muligt bør der anvendes forskellige tegn og typografiske midler. I ovennævnte artikelforslag er der gjort brug af en strukturindikator, idet alle fraseangivelser er skrevet med kursiv, således at de træder tydeligt frem i skriftbilledet. Dette har den yderligere fordel, at man ikke behøver at fremhæve parafraseangivelserne, da fraseangivelserne fungerer som direkte indgang til disse. For yderligere at gøre artiklen brugervenlig er det leksikografiske symbol ▲ anvendt som blikfang (strukturindikator), således at brugeren aldrig er i tvivl om hvor søgefeltet for flerordsforbindelser befinder sig. I denne artikel er der også inkorporeret grammatiske oplysninger. Der er i vinkelparenteser (strukturindikatorer) angivet, at de danske ækvivalenter i betydning 1 og 2 danner uregelmæssig pluralis (nulfleksion). Hvorvidt det er relevant med sådanne grammatiske oplysninger afhænger af målgruppens leksikografiske behov. I en ordbog, der benyttes til oversættelse eller tekstproduktion er det fordelagtigt at angive grammatiske uregelmæssigheder på målsproget. Det har især betydning i de tilfælde i hvilke lemmaet ikke er undergivet grammatiske uregelmæssigheder som tilfældet er i artiklen **contribution**.

I fraseangivelserne har jeg valgt at skrive lemmaet fuldt ud i stedet for at benytte et pladssymbol eller gentagelsestegn som fx en tilde. Dette optager ganske vist mere

plads, men brugeren er aldrig i tvivl om lemmaets form i en given frase. Et eksempel på, at lemmaet som subadresse i fraseangivelsen er forskelligt fra lemmaets form i lemma-positionen kan ses i følgende omarbejdede artikel:

contract in restraint of trade n [aftale om, at en person ikke må udøve et bestemt erhverv - kan være tids- eller områdebegrænset]

konkurrencebegrænsende aftale n

▲ *a contract in restraint of trade is prima facie illegal* en konkurrencebegrænsende aftale er almindeligvis uden retsvirkning; *contracts in restraint of trade are void unless it is proved that the restraint is reasonable* konkurrencebegrænsende aftaler er ugyldige medmindre det godtgøres, at konkurrenceforbudet er rimeligt

= konkurrenceforbud

→ non-competition clause; restrictive agreement

Fremgangsmåden med at benytte hele ordet i subadressen er specielt fordelagtig i de tilfælde i hvilke lemmaet har uregelmæssig fleksion (fleksionsmorfologiske uregelmæssigheder), fx det engelske verbum *catch*, der bøjes *caught* i datid. Endvidere er der anvendt semikolon efter parafraseangivelserne til at adskille disse fra de efterfølgende fraseangivelser. Semikolon anvendes som strukturindikator idet det angiver strukturen i den artikeldel, der behandler flerordsforbindelser og adskiller samtidig to mikrostrukturelle enheder. En væsentlig årsag til at anvende semikolon er, at man ofte kommer ud for, at fraseangivelser og parafraseangivelser indeholder almindeligt komma, og et komma ville derfor ikke kunne anvendes som strukturindikator. Når man vælger strukturindikatorer bør hver strukturindikator kun have ét formål for at undgå misforståelser, og alle strukturindikatorer bør være forklaret i ordbogens brugervejledning.

6. Andre mikrostrukturelle ovejelser

I forbindelse med alordbogsprincippet bør det endvidere være en af leksikografens primære opgaver at præsentere ordbogsbrugeren for en så nøjagtig og detaljeret leksikografisk behandling som muligt. Dette aspekt er dog afhængigt af flere faktorer, herunder ordbogens funktion, målgruppens leksikografiske behov samt ordbogens fag. I en engelsk-dansk juridisk ordbog for translatørstuderende må brugeren - hvis ikke andet er anført - kunne forvente at fagområdet engelsk ret (og dansk ret) behandles så grundigt, at unøjagtigheder undgås. I de valgte artikler fra FEO finder man en unøjagtighed i den leksikografiske behandling i artiklen **contractor**. Den unøjagtighed det drejer sig om i dette tilfælde er manglen på forskellige ækvivalenter af et polysemt lemma. I denne artikel er lemmaet behandlet som om det er monosemt selv om det faktisk er polysemt, eller i hvert fald svarer til flere danske juridiske udtryk. Der er ikke her tale om, at lem-

maet har én betydning inden for juridisk sprog, og en anden betydning uden for juridisk sprog, så den polyseme behandling er særdeles relevant i denne artikel. Inden for den juridiske sprogbrug svarer det engelske lemma i hvert fald til to forskellige danske juridiske ord, og dette burde have været angivet i artiklen. Et forslag til hvordan den i FEO unøjagtige behandling af lemmaet kunne forbedres fremgår af følgende omarbejdede artikelforslag:

contractor n [person, der har indgået en retligt bindende aftale]

1. **kontrahent** n [den, der deltager i og er bundet af en kontrakt]

= aftalepart; kontraktpart

2. **entreprenør** n [den, der i henhold til en entreprisekontrakt påtager sig at udføre bygge- eller anlægsarbejde på andens faste ejendom]

→ contract for services; independent contractor

3. **værksmester** n [ældre betegnelse for den, der mod vederlag forpligter sig til at frembringe et arbejdsresultat for en anden - omfatter fx entreprise og befragtning]

I den omarbejdede artikel bør man lægge mærke til, at der er angivet tre, og ikke to delbetydninger af lemmaet. Den tredje betydning er medtaget for fuldstændighedens skyld, men da det fremgår af den faglige forklaring i delbetydning 3, at den danske ækvivalent er en ældre betegnelse, kan den med fordel udelades. Det kan så undre, at FEO har valgt at angive en ukurant dansk ækvivalent i stedet for de mere kurante ækvivalenter, der er angivet i delbetydning 1 og 2. Det bør principielt være således, at kun kurante ækvivalenter angives, da disse er umiddelbart anvendelige, medmindre ordbogens erklærede formål er at give en diakron leksikografisk behandling af et givet fagomådes ordforråd. Under delbetydning 1 er der angivet to synonyme til ækvivalenten. Disse er anført i alfabetisk rækkefølge, men såfremt der foreligger en frekvensliste over et fags ordforråd, bør man overveje at anføre synonymerne i en rækkefølge, der afspejler denne frekvens. Uanset om man vælger at benytte det alfabetiske princip eller frekvensprincippet ved anførelse af mere end ét synonym, bør rækkefølgen altid være klart beskrevet i ordbogens brugervejledning. Til adskillelse af synonymerne under delbetydning 1 og til adskillelse af krydshenvisningerne under delbetydning 2 er der som strukturindikatorer anvendt semikolon. Dette skyldes, at semikolon allerede er anvendt til adskillelse af parafraseangivelser og fraseangivelser, og da det således fungerer som adskillelестegn mellem forskellige mikrostrukturelle enheder, kan man med fordel fortsætte med at anvende det på denne måde. Dette er en konsekvent fremgangsmåde, og den pågældende strukturindikator har kun én funktion.

Det bør dog bemærkes, at der i artiklerne fra FEO forekommer eksempler på en an-

givelse af flere danske ækvivalenter til ét engelsk lemma. Det bedste eksempel på dette er artiklen **contract**. I denne artikel angives tre danske juridiske udtryk som ækvivalenter til lemmaet. Inden for dansk juridisk sprogbrug anses disse tre ord ofte for at være synonyme udtryk, men der er dog en tendens til, at de ikke benyttes som fuldsynonymer. Eller sagt på en anden måde: i nogle tilfælde vil der være en tendens til at benytte det ene af de tre ord frem for de to andre (jf. *afbetalingskontrakt* vs. *konkurrencebegrænsende aftale*). For at gøre brugerne opmærksom på dette, er det nødvendigt at inkorporere oplysninger, der fortæller om dette. En omarbejdelse af denne artikel kan fx resultere i en ordbogsartikel med følgende indhold:

contract *n* [retligt bindende aftale, der stifter et retsforhold mellem parterne → §8]

1. **kontrakt** *n* [betegner almindeligvis en retligt bindende aftale]

▲ *a contract may be modified or terminated by the mere agreement of the parties* parterne kan ved aftale ændre en kontrakt eller bringe den til ophør; *at the time and place and in the form required by the contract* på det tidspunkt, på det sted og i den form kontrakten fastsætter; *the buyer may declare the contract avoided* køber kan hæve kontrakten; *the final wording of the contract* kontraktens endelige udformning; *the contract is up* kontrakten er udløbet; *keep the contract alive* fastholde kontrakten; *within the general scope of this contract* inden for nærværende kontrakts almindelige rammer; *contracts to be couched in a special form* formbundne kontrakter

≠ familieaftale

→ domestic agreement

2. **aftale** *n* [omfatter både retligt bindende og retligt uforpligtende aftaler]

▲ *at the time of the conclusion of the contract* på tidspunktet for kontraktens indgåelse; *the contract has been reduced into writing* aftalen foreligger i skriftlig form; *time is of the essence of the contract* tiden er af afgørende betydning for aftalen; *it is his job to procure or conclude contracts* det er hans job at udvirke eller indgå aftaler

→ agreement

3. **overenskomst** *n* [retligt bindende aftale, især kollektiv aftale mellem organisationer]

→ collective agreement

I denne omarbejdede artikel er der med polysemiindeks opnået en klar skelnen mellem de forskellige betydninger, og faglige forklaringer viser anvendelsen af de tre ækvivalenter. Herudover er der også et eksempel på en henvisning til en ordlisteekstern ordbogsdel i form af en eksplicit krydshenvisning: _ §8. Artiklen indeholder således en henvisning til en (imaginær) fagintroduktion i den omarbejdede ordbog, og et uddrag fra denne fagintroduktion kunne eventuelt have følgende udseende:

§7 Aftalers indgåelse

Engelsk og amerikansk aftaleret er baseret på overenskomstprincippet, ifølge hvilket et løfte eller tilbud først bliver bindende når det accepteres. Dansk aftaleret er baseret på løftet teorien, ifølge hvilken et løfte eller tilbud er bindende så snart det er kommet til modtagerens kundskab.

§8 Aftalers form

I engelsk og amerikansk ret sondres mellem to typer kontrakter. **Specialty contracts** er formbundne kontrakter, der oprettes skriftligt og forsynes med segl. **Simple contracts** er formløse kontrakter og kan være skriftlige eller mundtlige. Simple contracts opnår kun gyldighed såfremt **consideration** erlægges, hvorimod specialty contracts ikke kræver consideration, men opfyldelse af formkravene.

I dansk ret er hovedreglen, at kontraktens gyldighed ikke er betinget af bestemte formers iagttagelse.

§9 Aftalers gyldighed

.....

En sådan fagintroduktion kan være placeret før eller efter ordlisten, og den er af særlig betydning for brugere, der ikke er fagspecialister. Fordelen ved en fagintroduktion er, at man undgår at overbelaste artiklernes mikrostruktur idet man kan samle generelle oplysninger, som er relevante for flere lemmata - og dermed for flere artikler - et enkelt sted. Så i stedet for at skrive de samme oplysninger i flere artikler, kan man henvise fra artiklerne til fagintroduktionen. Såfremt fagintroduktionen er inddelt i nummererede paragraffer som i forslaget er det let at henvise til oplysningerne i denne.

Det er dog ikke gjort med dette, da det engelske udtryk **contract** bruges i flere betydninger end de her nævnte. En yderligere bearbejdning af artiklen resulterer i, at følgende oplysninger medtages:

4. **retshandel** n [bindende erklæring om stiftelse, forandring eller ophævelse af en rettighed]

▲ *any contract, dealing or transaction* enhver retshandel
→ act in law

5. **entreprise** n [aftale om udførelse af bygge- eller anlægsarbejde på en andens faste ejendom]

▲ *he was awarded a contract* han fik tildelt en entreprise;
he obtained a contract for a project han fik et projekt i
entreprise; *he won the contract for the work* han fik
arbejdet i entreprise

→ contract for services

6. **slutseddel** n [foreløbig aftale om overdragelse af ejendomsretten til fast ejendom - forkortelse for contract for the sale of land]

▲ *exchange of contracts* udveksling af slutsedler
→ contract for the sale of land

7. **kontraktforhold** n [betegner det samlede retsforhold mellem parterne, dvs. både kontrakten og den lovgivning og retspraksis, som har indflydelse på hele retsforholdet]

▲ *he has an action in both contract and tort* han kan
anlægge sag både i og uden for kontraktforhold

→ contractual relationship

Selv om artiklen således bliver forholdsvis omfattende, virker skriftbilledet ikke uoverskueligt, idet der er anvendt forskellige sekvenseringsenheder. Anvendelsen af strukturindikatorer, som fx fed skrift, kursiv skrift, skarpe parenteser og semikolon, samt det leksikografiske symbol ▲ betyder, at det er nemt at overskue artiklen og finde rundt i den. Det kan måske virke overvældende med alle disse oplysninger i de omarbejdede ordbogsartikler, men det er ikke altid nødvendigt med mange oplysninger. Eksempler på ordbogsartikler med et minimum af oplysninger er de såkaldte **krydshenvisningsartikler**. Af de udvalgte artikler fra FEO er der én, der med fordel kan ændres til en krydshenvisningsartikel, og dette er **contract to sell**. Som krydshenvisningsartikel kan den have følgende udformning:

contract to sell n = agreement to sell

I denne artikel henvises brugeren til lemmaet **agreement to sell**, i hvis artikel de relevante oplysninger vil være at finde. På denne måde undgår man at gentage oplysninger, og hvis det er muligt, kan man lade artiklen med det mest frekvente af to eller flere synonymer være den fuldstændige artikel.

7. Afslutning

Sammenfattende kan man sige, at der i tosprogede juridiske ordbøger for den fagligt ukyndige bruger skal være flere oplysninger end i FEO. Her er alordbogskonceptet vel-egnet. Eksempler på hvordan nogle artikler fra FEO kunne se ud efter en omarbejdning efter alordbogsprincippet og med behørig hensyntagen til de nødvendige oplysninger viser klart, at der kræves megen plads. En mulig løsning på dette problem er at begrænse den potentielle lemmabestand ved at udarbejde tosprogede fagordbøger, som behandler et snævert afgrænset fagområde, dvs. delfagsordbøger i stedet for hovedfagsordbøger. Dette medfører, at man med fordel kan begrænse sig til at udarbejde juridiske ordbøger inden for disciplinens delfag, fx en kontraktretlig ordbog, en familiretlig ordbog og en strafferetlig ordbog. På denne måde vil man kunne udarbejde ordbøger, der dækker mest muligt af ordforrådet inden for et sådant delfagsområde, og samtidig vil der blive plads til de meget omfattende oplysninger, som alordbogsprincippet kræver (jf. Lenstrup 1992 og Nielsen 1992). Da alordbogsprincippet indebærer en væsentlig belastning af den leksikografiske mikrostruktur, vil det være en stor fordel for brugeren, hvis man i alordbogens artikler gør udstrakt brug af strukturindikatorer som sekvenseringsenheder, således at den mikrostrukturelle sekvens klart fremgår af skriftbilledet. Dette medfører, at artiklen inddeles i klart afgrænsede søgefelter, således at brugeren hurtigt kan få tilgang til den type oplysninger han/hun søger.

Litteratur

- Basisordbog eng-da = Nielsen, Sandro: *Engelsk-dansk juridisk basisordbog*. København, Munksgaard 1993.
- Duvå, Grete/Laursen, Anna-Lise/Maidahl, Lisbet 1992: Brugerundersøgelse vedrørende oversættelse af fagtekster. I: Ruth Vatvedt Fjeld (red.): *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.-31. mai 1991*. Oslo: Nordisk forening for leksikografi: 105-133.
- FEO = Nielsen, Feodor: *Anglo/amerikansk-dansk retsordbog*. København: Nyt Nordisk Forlag Arnold Busk 1975.
- Grindsted, Annette 1988: *Principper for præsentation af ækvivalenter i oversættelsesordbøger*. Odense: Odense Universitetsforlag.
- Hausmann, Franz Josef 1989: Wörterbuchtypologie. I: F.J. Hausmann/O. Reichmann/H.E. Wiegand/L. Zgusta (red.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol. 1*. Berlin/New York: de Gruyter: 968-981.

- 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 (red.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol 1*. Berlin/New York: de Gruyter: 328-360.
- Kromann, Hans-Peder/Riiber, Theis/Rosbach, Paul 1984: Überlegungen zu Grundfragen der zweisprachigen Lexikographie. I: Herbert Ernst Wiegand (red.): *Studien zur neuhochdeutschen Lexikographie V*. Hildesheim/New York: Olms: 159-238.
- Lenstrup, Rita 1992: Den selektive specialordbog. I: Ruth Vatvedt Fjeld (red.): *Nordiske Studier i leksikografi. Rapport fra Konferanse om leksikografi i Noden 28.-31. mai 1991*. Oslo: Nordisk forening for leksikografi: 478-482.
- Malkiel, Yakov 1975: A Typological Classification of Dictionaries on the Basis of Distinctive Features. I: Fred W. Householder/Sol Saporta (red.): *Problems in Lexicography*. Bloomington: Indiana University: 3-24.
- Nielsen, Sandro 1990: Contrastive Description of Dictionaries Covering LSP Communication. I: *Fachsprache/International Journal of LSP 3-4*: 129-136.
- Nielsen, Sandro 1992: Forholdet mellem enkeltfagsordbøger og alordbøger. I: R. Vatvedt Fjeld (red.): *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.-31. mai 1991*. Oslo: Nordisk forening for leksikografi: 275-287.
- Wiegand, Herbert Ernst 1988: Was eigentlich ist Fachlexikographie? Mit Hinweisen zum Verhältnis von sprachligem und enzyklopädischem Wissen. I: Horst Haider Munske/Peter von Polenz/Oskar Reichmann/Reiner Hildebrandt (red.): *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 (red.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie, vol. 1*. Berlin/New York: de Gruyter: 409-462.
- Zgusta, Ladislav 1971: *Manual of Lexicography*. Haag/Paris: Mouton.

Vart tog Svensson vägen?

Om ordböcker och proprier

Kristina Nikula

Namn eller proprier (förnamn, ortnamn etc.) hör inte hemma i allmänspråkliga ordböcker, då det begränsade utrymme som står till buds borde nyttjas för information som ordboksanvändarna blivit vana att hitta och som därmed efterfrågas oftare än namns stavning, uttal och böjning. Dessutom kräver namn en annan behandling än det "normala" ordförrådet.¹ Namn saknar visserligen inte semantisk betydelse, men då denna är ytterst begränsad och ett namn ju inte kan användas korrekt utgående från ordboksdefinitioner, utan först när man vet vem som heter vad, är det heller ingenting som talar för att ett namn skall tas med i en definitionsordbok. – Varje gång det blir aktuellt att stava till *Helsingborg* eller uttala *Jokkmokk* känns det emellertid som en brist att namn i större utsträckning inte finns i de vanliga ordböckerna. Som en direkt miss från ordboksförfattarens sida upplever jag att namn som *Svensson* och *Sörgården* o.a., som kan misstänkas för att ha också någon annan än en rent onomastisk betydelse, inte finns med.

1. Syfte

Substantiven har av gammalt delats in i appellativer och proprier och man har med hjälp av formella och semantiska kriterier försökt dra en gräns mellan appellativerna och proprierna utan att för den skull vara särdeles framgångsrik, då det knappast existerar någon klar gräns mellan dessa två kategorier av ord.² Om man inom kategorin proprier ytterligare skall försöka avskilja lexem som är levande som namn och samtidigt åtföljs t.ex. av starka konnotationer, utan att för den skull kunna betraktas som lexikaliserade, kan uppgiften förefalla övermäktig. – Vad det är som gör proprierna speciella, vilka som hör hemma i ordböckerna och vilka egenskaper hos dessa som bör framhåvas kommer jag att diskutera i detta bidrag. Den slutliga utformningen av ordbeskrivningen överlåter jag däremot på den praktiskt arbetande lexikografen.

2. Kännetecken på proprierna

2.1 *Skrivsättet*

I debatten om införandet av husläkare i Sverige konstateras i en artikel följande: "Frågan är: blir det ett jollrande medicinskt Sörgården av hela Sverige?"³ Det vore för lexikogra-

fen onekligen enklast att som namn betrakta de språkliga tecken som skrivs med stor bokstav – redan i detta exempel skulle två uppslagsord inbesparas. Så länge man håller sig till modern skriven text på svenska klarar man sig antagligen rätt långt på så vis, men redan i fråga om tal är detta verktyg odugligt. Det samma gäller för språk där alla substantiv skrivs med stor bokstav. Att ett tecken skrivs med stor begynnelsebokstav är ju inte orsak till utan följderna av att detta avses utgöra eller uppfattats som ett namn. Det vanskliga i att låta skrivsättet vara avgörande för kategoritillhörigheten framgår inte minst av Språknämndens rekommendationer för användningen av stor bokstav i svenskan.⁴

2.2 Böjningen

Som kännetecken på proprierna brukar anges att de inte kan böjas i pluralis och att de inte kan uppträda i obestämd form eller anta bestämd form, då de redan i sig är bestämda. Ytterligare brukar anges att de inte heller kan kombineras med restriktiva bisatser. Åtskilliga exempel visar dock att dessa kriterier är ohållbara, även om artikellös form kan anses vara den omarkerade formen eller normalformen för ett proprium.⁵ Men inte ens om detta kännetecken är proprietens ensamt. Också appellativer kan i vissa situationer uppträda med nollartikel, t.ex. i tidningsrubriker och på skyltar (*Dynamik* i lågkonjunkturens spår⁶; *Infart, Garage, Återvändsgränd*).

2.3 Betydelsen

Man är tämligen överens om att proprierna refererar till **en** unik referent. Däremot är åsikterna om propriernas betydelse, något förenklat uttryckt, ofta diametralt motsatta. Förespråkarna för den ena ståndpunkten anser att proprierna saknar betydelse, medan förespråkarna för den andra sidan anser proprierna vara de mest betydelsrika substantiven.⁷

3. Textlingvistisk beskrivning av proprierna

Ett försök att närma sig namnproblematiken ur en ny synvinkel har gjorts inom textlingvistiken, där man utgår ifrån att egenskapen att utgöra ett proprium inte är inherent hos ett lexem, utan att denna egenskap tillförs detta utifrån den text där lexemet ingår. Ett språkligt tecken har sålunda inte en naturgiven position som proprium, utan denna status har historiskt sett utvecklats (förnamn, efternamn, ortnamn) eller kan utvecklas med hjälp av textuella signaler, transformationssignaler, eller pragmatiska determinanter. Då tecknet som sådant sålunda inte kan fungera gränsöverskridande, utan proprialiseringen resp. appellativiseringen initieras utifrån, innebär det att tecknet med dess formella och/eller semantiska egenskaper inte, som tidigare har skett, kan studeras isolerat, utan att detta måste ses som en beståndsdel av en text och vidare att proprium som kategori uppstår först inom ramen för text. Genom att namn är förknippade med specifika sociala situationer (namngivning, kontakt med myndigheter etc.) medvetandegörs de ständigt som en exklusiv kategori och medlemmarna i en språkgemenskap är därigenom

om vanligen också medvetna om vilka metaspråkliga signaler som tjänar till att urskilja namn.⁸

3.1 Transformationssignaler

De textuella signaler som är med om att skapa ett proprium är av skilda slag: de kan vara textgrammatiska, -semantiska, fonologiska, fonetiska och grafiska (t.ex. stor bokstav). Speciella textsignaler är emellertid inte alltid av nöden för att ett lexem skall uppfattas som ett proprium. Detta gäller först och främst förnamnen, vars status som egennamn kodifierats inom språkgemenskapen och hos vilka denna tack vare konvention, tradition och frekvens är inherent. Denna egenskap försätter förnamnen i den unika ställningen att de själva också kan fungera som namnskapande signaler och utpeka ett följande lexem som t.ex. efternamn. De enskilda språken tillhandahåller i regel en särskild namnkorpus i vilken förnamn, namn på länder och städer o.d. ingår.⁹

Av de textuella signaler som är med om att skapa ett proprium kommer av utrymmes-skäl bara namngivningsakten och de signaler som är förknippade med denna att behandlas i detta sammanhang. Denna akt, som på sätt och vis upprepas varje gång man omtalar namnet på en person, kommer inte alltid till uttryck med hjälp av performativa uttryck av typen *jag döper dig till...* utan kan lika väl markeras med uttryck som *jag heter...*, *mitt namn är...*, *vi kallar det...* etc.¹⁰ En primär namngivningsakt, i den medvetna avsikten att skapa ett namn kan också saknas och lexemet har i sådana fall genom användningen småningom utvecklats till och börjat användas som namn. Vid namngivning med fullständigt lexikaliserade proprier, i allmänhet förnamn, skapas naturligtvis inte namnet genom namngivningsakten, det endast ges, vilket just är ett uttryck för lexikaliseringsgraden. Vid lägre lexikaliseringsgrad "återskapas" proprier.

Ur textlingvistisk synvinkel kan ett proprium sålunda definieras exempelvis på följande sätt: "Ein Sprachzeichen (bzw. eine Sprachzeichenkette) fungiert dann als Eigenname, wenn es als ein solcher intendiert [...] und über geeignete kontextuelle Sprachmittel wie auch situative (pragmatische) Signale als ein Name verstanden ist [...]".¹¹ Om den mer eller mindre medvetna avsikten ses som avgörande för om ett språkligt tecken skall betraktas som ett proprium eller inte, säger det sig självt, att detta inte kan betraktas enbart ur grammatisk och/eller semantisk synvinkel, utan man måste fråga sig vilka språkliga och utomspråkliga medel som tillsammans stagna bidrar till att vissa språkliga tecken uppfattas som namn.

4. Extension och intension

I ett försök att sammanjämka åsikterna om förekomsten resp. frånvaron av betydelse hos namn inför Kalverkämper¹² motsatsförhållandet extension-intension i diskussionen. Eftersom extensionen hos proprier är begränsad till specifika namnbärare får namnet enligt Kalverkämper ett desto större intensionsdjup. I sin uppfattning av intensionen företräder denne sålunda på det hela taget samma åsikter som inom nordisk namnforskning framförts av bl.a. Hedquist och Pamp¹³. Extensionen utgör inte heller något pro-

blem, men det är däremot inte utan svårighet begreppet intension låter sig utnyttjas i frågan om namn¹⁴, eftersom denna i och med den minimala extensionen blir så innehållsrik (innehåller all tänkbar kunskap om namnbäraren)¹⁵ att den i praktiken inte går att beskriva med språkvetenskapliga medel. Även om Kalverkämper sålunda i princip anser det som förfelat att definiera namns semantiska innehåll utifrån intensionen, företräder han ändå i sina nyaste arbeten¹⁶ åsikten att intensionen bör beaktas.

Kalverkämpers användning av begreppet intension kräver en kommentar. Till sin extension kan ett namn som *Lisa_J* anses vara maximalt begränsat, då det omfattar endast den bestämda *Lisa_J* med vilken den talande och åhöraren förknippar namnet i en bestämd situation. Eftersom intension och extension normalt står i omvänt förhållande till varandra, skulle logiskt följa att *Lisa_J* har en maximalt omfattande intension, dvs. en intension som kännetecknas av ett ytterst stort antal semantiska särdrag. Men trots att man känner alla särdrag, trots att man vet "allt" om *Lisa*, vet man inte att hon heter *Lisa* om man inte vet att just hon kallas så. Enligt min mening förefaller alltså den relation mellan intension och extension som gäller för appellativer vara upphävd hos *proprier*. Det enda verkligt väsentliga är "namngivningsakten", vilken sålunda på sätt och vis övertar den funktion intensionen har hos appellativerna.

Jag har i detta fall utgått från antagandet att *Lisa* varje gång namnet påträffas hos en ny namnbärare utgör ett nytt namn homonymt med alla andra namn *Lisa*. Saken förändras emellertid inte, även om man ur alla namn *Lisa_{1-n}* abstraherar ett slags "över-*Lisa*". Visserligen kommer extensionen att bli åtskilligt mycket större och omfatta alla som heter *Lisa*, men intensionen kan knappast omfatta så många andra särdrag än /+kvinnlig./ För namn gäller sålunda att extensionen är given genom namngivning i varje enskilt fall. De få semantiska (språkliga) särdrag som föreligger, t.ex. /+kvinnlig/ i fråga om *Lisa*, styr sålunda primärt namngivningsakten och endast i ytterst ringa mån identifieringen av referenten. Detta innebär att en ny aspekt införs i diskussionen av relationerna mellan encyklopedisk och lexikalisk kunskap. Själv ser jag lexikalisk kunskap som sådan encyklopedisk kunskap som med nödvändighet är förknippad med en lexikalisk enhet för att denna i olika tänkbara kontexter skall kunna fylla sin teckenfunktion.

Diskussionen om betydelsen hos namn försvåras ofta av bristen på definitioner på vad som egentligen avses med betydelse. Det förefaller som om t.ex. Kalverkämper med betydelse avser all tänkbar kunskap som kan härledas om en namnbärare – konnotationerna eller "das atmosphärische Umfeld" utesluter han dock, då dessa i stor utsträckning är beroende av utomspråkliga faktorer¹⁷. Kunskapen om namnbäraren kan sålunda vara omfattande och den växer därtill i takt med att man bättre lär känna namnbäraren (viss kunskap kan för all del också falla i glömska). Om däremot med betydelse avses lexikalisk betydelse eller den kunskap om referenten som är av betydelse för beteckningsfunktionen¹⁸ kan betydelsen hos t.ex. *Lisa* komprimeras till 'namn på feminin person', där just *namn* anger nödvändigheten av en namngivningsakt för att det skall vara möjligt att individualisera och identifiera namnbäraren och att namnet sålunda måste läras in från fall till fall.

5. Namngivningsakten

I fråga om propriier träder sålunda som jag valt att se saken i stället för intensionen namngivningsakten, då det är denna som avgör vem som heter t.ex. *Lisa* och som följaktligen ger namnet dess extension. Att själva namngivningen eller informationen om en persons namn får ersätta intensionen innebär ett annat synsätt än t.ex. Kalverkämpers. Dennes försök att bestämma extensionen för ett proprium med hjälp av den textuella och/eller pragmatiska situationen och att därefter utfylla intensionen med hjälp av texten och kommunikationssituationen försätter propriiet i liknande situation som pronomen, då ju också dessa får sin betydelse inom ramen för texten. Kontextinformation kan visserligen väsentligt underlätta identifikationen av namnbäraren, men den ersätter inte namngivningsakten.

6. Proprier med sekundär appellativisk funktion

Det är inte ovanligt att proprier i synnerhet i vissa typer av text uppträder föregångna av obestämd artikel (ex. 1-5).

1. Det var en förstörelse som Karelen aldrig repade sig ifrån, *ett Karelen* som under medeltiden i löst förbund med Novgorod fått uppleva långa perioder av välstånd och vitalitet.¹⁹

Artikelvändningen brukar i fall som detta tolkas som ett tecken på att propriiet bytt kategori och utgör ett appellativ²⁰. I likhet med ett lexem som i text får sin funktion som proprium med stöd av olika kontextuella signaler, kan ett proprium sekundärt ges appellativisk funktion med hjälp av andra typer av textsignaler. För att en förändring av funktionen skall vara möjlig krävs emellertid att vissa villkor är uppfyllda; Kalverkämpers räknar ursprungligen med två, men i ett av sina nyaste arbeten med tre olika villkor²¹: 1. namnet skall som namn vara bekant för kommunikationsparterna, 2. transponeringen är avhängig av kontexten och i synnerhet av att signaler (*en, ett*) anger den nya användningen och 3. propriets funktion skall efter transponeringen vara en annan än den lexemet har i sin egenskap av namn. – Den obestämda artikeln *en, ett* signalerar i kombination med ett proprium inte att förhandskunskapen om referenten här och nu inte längre gäller eller att propriiet nödvändigtvis skall uppfattas som appellativ²², utan snarast att vissa egenskaper hos referenten fokuseras (ex. 1) eller att namnet uppträder i metaforisk betydelse närmast i form av en jämförelse. Ytterligare kan på detta sätt anges att osäkerhet råder i fråga om vem namnet refererar till (ex. 2). – *Ett Karelen* i exempel (1) måste alltså uppfattas som ett proprium, men övertar i kontext appellativa funktioner.

Osäkerhet i fråga om referenten innebär det samma som talar- resp. hörarobestämdhet²³, dvs. den talande markerar med den obestämda artikeln, första gången han omtalar en person, att han själv inte vet vem referenten är eller att han utgår ifrån att referenten är obekant för åhöraren.

2. Det ringde *en/nån' Andersson* för en stund sedan.

Obestämd artikel kan förutom osäkerhet i fråga om referenten också markera att det saknar betydelse vem den åsyftade egentligen är. Av exempel (3), som handlar om judarnas ställning i Granada, framgår att dessa nådde de högsta positioner i samhälle; vilka de enskilda positionsinnehavarna var, är kanske inte närmare känt eller av mindre intresse i sammanhanget.

3. Snart hittar vi halvöns forna parias skrida fram i siden och raffiga turbaner, med tillgång till rikets högsta poster – *en Samuel* var storvesir i 28 år – och bländande världen också med sin lärdom.²⁴

Det är sålunda inte enbart namnbärarna som på normalt sätt identifieras och aktualiseras med namnet, utan (vissa) egenskaper eller prestationer som förknippas med dessa och som gör dem enastående i sitt slag lyfts på detta sätt fram och namnbärarna används därigenom som exempel på eller prototypen för personer av en viss typ. För den initierade läsaren definieras personen i fråga med hjälp av sin stereotyp. Även om propriet föregås av artikel är detta fortfarande primärt ett proprium, eftersom identifikationsprocessen först måste äga rum för att kunskapen om namnbäraren skall kunna aktualiseras och därefter utnyttjas i olika syften. Anders Chydenius i exempel (4), som egentligen var präst, men också riksdagsman, är framför allt känd som en skicklig nationalekonom. I vilken egenskap han främst anförs är inte helt entydigt, men det förefaller av kontexten att döma vara som en framstegsvänlig företrädare för prästgårdskulturen.

4. Eller tag *en Anders Chydenius*, inte nog med att han vid riksdagen i Stockholm med lysande framgång slogs för bondeseglationens frihet, som kyrkoherde i Gamlakarleby höll han [...] ²⁵

Om man riktar sig till en okänd publik är inte ovanligt att författarna för säkerhets skull antyder inom vilket område en namnbärare främst gjort sig känd. Detta framgår tydligt av exempel (4-5), där presentationen dels föregår namnen (ex. 5, *intellektuella*), dels följer efter (ex. 4, *kyrkoherde*). Presentationen kan också stå på båda sidor om namnet på en och samma gång²⁶. I fråga om unireferentiella (*Hitler*, *Stalin*) namn är en presentation onödig.

5. Bland de intellektuella vann Hitler relativt få proselyter, hur gärna än *en Tito Colliander* eller *Örnulf Tigerstedt* lät sig inviteras till Weimar 1942.²⁷

Jag har av utrymmesskäl här behandlat bara ett fåtal vanligen mera sällan uppmärksammade betydelser hos egennamn som föregås av obestämd artikel. I regel brukar konstateras att artikel och proprium (familjenamn) anger medlem av familjen eller innebär en jämförelse med den egentlige namnbäraren.

7. Svensson och Sörgården

Det är inte i första hand de lexem som på ett eller annat sätt igenkänns som typiska namn som är intressanta med tanke på en allmänspråklig ordbok. Inte heller är de som appel-

lativer lexikaliserade namnen något problem, i synnerhet om ursprunget fallit i glömska (*McAdam - makadam* o.d.), utan främst de namn som intar ett slags mellanposition och som oberoende av om faktabaserad information om namnbäraren redan lämnats i texten eller inte, som Kalverkämper²⁸ uttrycker saken, omges av "das atmosphärische Umfeld", also das nicht ganz greifbar mitschwingende Wissen: die Konnotationen". Konnotationerna är ju för det mesta mer eller mindre individuella beroende på ålder, erfarenheter, kulturell bakgrund o.d. hos språkbrukarna, men utöver dessa existerar också konnotationer i form av ett slags kollektiv erfarenhet eller traderade klischéer, som finns komprimerade i namn, som härigenom kommer att utgöra ett slags "nyckelnamn" med stort socialt konnotationsvärde. För att ett namn skall kunna anses utgöra ett nyckelnamn, krävs det att det återkommande uppträder i denna funktion, medan av en stark kontext motiverade ad hoc-konnotationer hos namnet inte gör detta till ett nyckelnamn. Ett för finlandssvenskan typiskt sådant nyckelnamn är *Långa bron*²⁹. I det inledande exemplet utgör i mitt tycke *Sörgården* också ett sådant namn, då väl detta för många innebär sol och sommar, röda stugor med vita knutar och flaggan i topp, dvs. kort sagt en idyll.

Med hänsyn till att långtifrån ens alla infödda språkbrukare, trots en långtgående consensus om vissa namns konnotationsvärde, alltid är på det klara med dessa nyckelnamns "betydelse", finner jag det angeläget att namn av den här typen med åtföljande betydelsebeskrivning bereds plats i ordböckerna, även om kommunikationen inte är dömd att direkt misslyckas om den ena kommunikationsparten inte är medveten om konnotationerna. Det är snarast så att kommunikationen för den okunnige så att säga kommer att föras på ett ytligare plan. Det dagliga umgänget med svenskstuderande som inte har svenska som modersmål visar också tydligt på behovet av ordboksinformation. Det kan visserligen hävdas att de flesta enspråkigt svenska ordböcker i huvudsak utarbetats för en målgrupp med svenska som modersmål, men ser man till vilka användarna i själva verket är, framstår gruppen i språkligt hänseende som synnerligen heterogen. Tilltagande internationalisering, invandring osv. kommer sannolikt också att högst väsentligt öka behovet av information.

Rent konkret innebär detta att ett namn som *Sörgården* i betydelse 'idyll' borde ges plats i ordböckerna vid sidan av *sörgårds-/Sörgårdsidyll* (SOB). Också *svensson/Svensson* i betydelsen 'svensk' borde finnas med och inte som nu enbart *medel-/Medelsvensson* (SOB). Till den här kategorin hör förmodligen också *Östermalm* eller ännu hellre *Öfre Östermalm*? I fall som dessa visar det sig också hur missvisande skrivsättet kan vara som tecken på ett lexems status som namn, då ju svenskan i många fall tillåter skrivning med såväl stor som liten begynnelsebokstav och man med enbart detta som kriterium sålunda kan förbigå namn som av flera orsaker borde förekomma som uppslagsord.

Sammanfattning

Det har gällt som en oskriven lag att namn inte ingår i allmänspråkliga ordböcker, vilket sammanhänger med propriernas speciella egenheter i synnerhet i fråga om betydelsen.

I detta bidrag har jag diskuterat dessa egenheter och kommit fram till att den för appellativer typiska omvända relationen mellan intension och extension är upphävd och att namngivningsakten nästan helt ersätter intensionens funktion hos proprierna. Däremot har jag inte berört t.ex. den textlingvistiskt viktiga men lexikologiskt/lexikografisk mera sekundära motivbetydelsen (genomskinligheten) hos många proprier (*Kyrkbacken*, motivbetydelse 'backen upp mot kyrkan'). Alla proprier kan kontextuellt användas med appellativ funktion utan att deras status som proprier därmed upphävs. Viktiga att ta med i allmänspråkliga ordböcker är enbart sådana proprier vars användning appellativt i någon bestämd innebörd konventionaliserats, t.ex. *Svensson* med innebörden 'den typiske svensken'.

Noter

- 1 Falck-Källquist 1993.
- 2 Fleischer 1992:4,23. Uppsatsen baserar sig på ett föredrag från 1963. Rostvik 1969:118.
- 3 Unge 1993:4.
- 4 Svenska skrivregler 1991:47 ff.
- 5 Nikula 1993b: 168f.
- 6 Vasabladet 28.7.1993.
- 7 Nikula 1993a: 171ff., 1993b: 171ff.
- 8 Kalverkämper 1994a: avsn. 3, 1994b: avsn. 3.
- 9 Kalverkämper 1994a: 4.1.3.
- 10 Jfr Kalverkämper 1978:129 ff, 1994a: 4.1.1.
- 11 Kalverkämper 1994a: avsn. 2, jfr 1978:386.
- 12 Kalverkämper 1978:62 ff.
- 13 Hedquist 1984:130, Pamp 1976:335.
- 14 Jfr Aschenberg 1991:23f.
- 15 Kalverkämper 1978:65.
- 16 Kalverkämper 1994c: 2.1.
- 17 Kalverkämper 1978:89
- 18 Nikula 1993a: 171 f., 1993b: 171 f.
- 19 Salminen 1988:54.
- 20 Thorell 1973:§115.
- 21 Kalverkämper 1978:327 resp. 1994a: 4.2.
- 22 Jfr Kalverkämper 1994a: 4.2.1.
- 23 Kolde 1992:147, Nikula 1993b: 170f.
- 24 Salminen 1988:23.
- 25 Salminen 1988:50.
- 26 Kalverkämper 1994a: 5.2.
- 27 Salminen 1988:73.
- 28 Kalverkämper 1994a: 5.1.
- 29 Nikula 1993b: 177ff.

Litteratur

- Aschenberg, H. (1991), *Eigennamen im Kinderbuch. Eine textlinguistische Studie*. Gunter Narr Verlag.
- Falck-Källquist, B. (1993), *On Proper Names in Historical-Etymological Dictionaries - with special reference to personal names*. (Manuskript).
- Fleischer, W. (1992), *Name und Text. Ausgewählte Studien zur Onomastik und Stilistik*. Max Niemeyer Verlag.
- Hedquist, R. (1986), *Eigennamens semantik*. I: *Florilegium Nordicum*. Acta universitatis Umensis. Umeå Studies in the Humanities 61, s. 127-134. Almqvist & Wiksell International.
- Kalverkämper, H. (1978), *Textlinguistik der Eigennamen*. Klett-Cotta.
- Kalverkämper H. (1994a), *Eigennamen in Texten*. Utkommer i Canisius, P., Tschauder, G., Herbermann, C-P. (Hrsg.), *Text und Grammatik*. Festschrift für Roland Harweg zum 60. Geburtstag.
- Kalverkämper, H. (1994b), *Zur Textgrammatik und Textsemantik der Eigennamen*. Utkommer iu Eichler, E. et al. (Hrsg.) *Namenforschung - Proper Name Studies - Les noms propres*. Ein internationales Handbuch zur allgemeinen und europäischen Onomastik. de Gruyter.
- Kalverkämper, H. (1994c), *Namenübersetzung*. Utkommer i samma som Kalverkämper (1994b).
- Kolde, G. (1992), *Zur Referenzsemantik von Eigennamen im gegenwärtigen Deutschen*. I: *Deutsche Sprache*. Zeitschrift für Theorie, Praxis, Dokumentation 2, s. 139-152. Erich Schmidt Verlag.
- Nikula, K. (1993a), *Produktnamn som lexikografiskt problem*. I: Fackspråk och översättningsteori. VAKKI-symposium XIII. Våren 13-14 febr. 1993, s. 167-179, Vasa..
- Nikula, K. (1993b), *Eigennamen och den internt tvåspråkiga ordboken*. I: *Svenskan i Finland 2*, s. 167-185. Meddelanden från institutionen för nordiska språk vid Jyväskylä universitet 9, Jyväskylä.
- Pamp, P. (1976), *Vad betyder Gösta Holm egentligen?* I: *Nordiska studier i filologi och lingvistik*, s. 330-341. Studentlitteratur.
- Rostvik, A. (1969), *Om klassificering av ortnamn*. I: *Namn och Bygd* 57, s. 112-132.
- Salminen, J. (1988), *Minnet av Alexandria*. Söderströms.
- SOB = Svensk ordbok (1986), *Språkdata*. Göteborgs universitet. Esselte Studium.
- Svenska skrivregler (1991), *Svenska språknämnden*. Almqvist & Wiksell.
- Thorell, O. (1973), *Svensk grammatik*. Esselte Studium.
- Unge, I. (1993), *Husläkare. Personlig omsorg eller vård på löpande band?* I: *Vi* 30, s. 4-5. Vasabladet 28.7.1993.

Kontrastiv metaforik

- presentation av ett lexikologiskt pilotprojekt

Kerstin Norén

En förstudie till projektet "Kontrastiv metaforik" ingår som en del i det s.k. lexikologi-programmet som humanistiska fakulteten vid Göteborgs universitet driver under en femårsperiod. Elzbieta Dura, Maja Lindfors-Viklund och Lena Rogström har, förutom jag själv, bidragit till projektet såväl med teoretiska synpunkter som med praktiskt arbete.

Först något om projektets namn. Om man med termen *kontrastiv* avser en jämförelse mellan två språk, så är namnet i viss mån missvisande. Visserligen har vi hittills mest gjort jämförelser mellan två språk, svenska och engelska, men avsikten är att fler språk så småningom ska behandlas. En mer korrekt term lär då ska vara *tvärspråklig*. Också termen *metaforik* är egentligen för begränsad. Ett syfte med projektet är att undersöka lexikaliska polysemier i olika språk, däribland metaforer som är så etablerade och förbleknade att de förtecknas i lexikon. Men också andra polysema betydelser är aktuella om de är etablerade. En mer adekvat term än metaforik vore alltså *polysemi*. Om rubrikers enda syfte vore att vara maximalt informationsrika borde projektet alltså kallas "Tvärspråklig lexikalisk polysemi". Men jag anser att rubriker också har till uppgift att vara kortfattade och låta bra, så att de därigenom blir lättare att komma ihåg och att referera till. Därav det valda projektnamnet.

Som underlag för analyserna har vi konsulterat stora enspråkiga definitionsordböcker. För engelskans del har vi huvudsakligen använt Longman Dictionary of Contemporary English och för svenskans del Svensk ordbok. Användningen av olika ordböcker har i sig medfört en del problem. Ett av dem rör själva uppdelningen i polysema betydelsemoment, som ju kan göras på lite olika sätt. Vi har dock haft god nytta av att de båda ovannämnda ordböckerna i stort sett genomgående ger exempel på varje polysem betydelse som anges. Ett annat problem rör fullständigheten i angivelserna av antalet olika betydelsemoment. Då man jämför två språks lexikon framgår snart att vissa av de olikheter man upptäcker snarare beror på lexikografiska än på språkliga skillnader, t.ex. genom att den ena ordboken tar upp ett betydelsemoment som förtigs i den andra. Förtigandet kan bero på skillnader i frekvens, etableringsgrad etc. men också på rena förbi-seenden. Ordböckernas uppgifter måste därför kompletteras av inhemska informanter, som kan ge en mer nyanserad bild än lexikon av polysema betydelsernas användbarhet och som kan fylla i sådana luckor som inte beror på en verklig skillnad mellan språkens lexikala polysemi.

En annan effekt av att konfronteras med olika ordböckers sätt att hantera lexikalisk polysemi är att man måste aktualisera frågan om vad lexikaliskt etablerad betydelse egentligen är. Vissa metaforiska betydelser kan exempelvis anges i ordböcker. Men hur skiljer sig dessa betydelser från metaforisk användning av ord i liknelser? En tumregel kan vara att lexikaliskt etablerade substantiviska metaforer kan användas självständigt, utan något jämförelseled uttryckt. Både *lejon* och *gasell* kan användas om människor i liknelser (med eller utan *som* insatt), t.ex. *han var (som) ett lejon*, *hon är (som) en gasell*, men bara *lejon* kan användas fritt om människor utan jämförelseled: *de unga lejonerna i af-färsvärlden*, **gasellerna dansade på verandan*. Alltså kan *lejon* betraktas som lexikaliskt polysemt, men inte *gasell*. Så är fallet i Svensk ordbok. För verbpolysemis del kan en tumregel vara att ett lexikaliskt etablerat betydelsemoment kan användas absolut, dvs. utan objekt (eller adverbial) som indikerar att det är fråga om metaforisk användning. Så kan t.ex. *du fattar* betyda 'du förstår' (minst) lika väl som 'du tar tag'. Om den mentala betydelsen av *fatta* inte vore etablerad skulle man alltid behöva säga *du fattar min idé* e.d. För ett verb som *nå* gäller i viss mån detta omvända förhållande. I sin kärnbetydelse kan verbet användas absolut: *jag når (inte)*, men metaforiskt är en sådan användning knappast tänkbar, utan det metaforiska objektet bör finnas med: *nå goda resultat*. Liknande resonemang kan föras för adjektiv och vissa adverb.

Användning av ord i ordspråk faller normalt också utanför den lexikaliska polysemis domäner. *Mycket skrik för lite ull* kan parafrastras ungefär med 'mycket rabalder / oväsen / skrik för litet resultat'. Det ser av ordspråket ut som om *ull* kan betyda 'resultat', men det betyder ändå inte att *ull* lexikaliskt har den polysema betydelsen 'resultat'. Man kan inte säga exempelvis *det blev ingen ull av förhandlingarna den här gången heller*. Ordspråkens ord refererar propositionellt till hela, likartade situationer och kan inte ges en betydelse var för sig. Det kan inte heller ord i idiom. Metaforisk användning av ord i exempelvis allusioner och allegorier är mestadels också tillfällig och anges inte i lexikon.

Att tumregeln för lexikalisk metaforik handlar om jämförelseled är ingen tillfällighet. Vår semantiska analysmodell är inriktad på (informella) betydelsekomponenter, och metaforik kan i en sådan modell beskrivas som utbyte av en, eller ofta flera, komponenter. Utbytet grundas på likhetsrelationer av något slag, vilket förutsätter just jämförelser. Andra förändringar bland ett ords betydelsekomponenter kan vara tillägg av komponenter. Ett exempel är när *rulle* 'cylinder' också kan få den mer specialiserade betydelsen 'film'. "Tillägget" kan också innebära ett val av endera värdet på någon variabel som är obestämd i utgångsbetydelsen. Att både tillägg och val behandlas som en kategori beror på att de båda innebär en precisering av betydelsen. Förlust av komponenter är också vanligt. Då blir betydelsen mer generell: *brud* 'kvinna vid giftermålet; vanligen ung och tilldragande' blir 'ung och tilldragande kvinna'. Det sistnämnda exemplet visar också en fjärde typ av förändring, som kan kallas omhierarkisering av komponenter - 'ung' och 'tilldragande', som inte är framträdande i kärnbetydelsen, blir viktigare i den utvidgade betydelsen. Inte sällan samverkar de fyra förändringstyperna med varandra vid bildandet av polysema betydelser. Metonymier kan t.ex. räknas som ett slags kombination av tillägg och omhierarkisering - nya komponenter i företeelser i närheten av kärnordet

lyfts fram. Detta förekommer i svenskan bl.a. när verbalsubstantiv som betecknar en verksamhet, t.ex. *administration*, också kan användas om lokalerna där verksamheten sker eller om personerna som utför verksamheten i fråga. Slutligen kan också neutralisering av olika värden förekomma vid polysemier.

Uttrycksmässiga tumregler för att avgöra om polysemier som bygger på tillägg, förlust eller omhierarkisering av en kärnbetydelses komponenter har blivit etablerade i lexikonet är svårare att hitta än tumregler för utbyte. Generellt gäller emellertid för många polysema betydelsemoment i förhållande till kärnbetydelsen att de, som en effekt av betydelseförändringarna, har en annorlunda syntaktisk ram: *?min farbror* (om icke-släkting), *?såga fiol* (jfr *såga ved*), *en dödare tavla* (*har jag sällan sett*) (jfr *?en dödare man...*).

Syftet med projektet antyddes redan inledningsvis. Det är fråga om en detaljerad kartläggning av dels vilka lexikaliska polysemirelationer som finns i mänskliga språk i allmänhet, dels mellan hur olika språks lexikon utnyttjar olika typer av polysemi samt vilka enskilda ord som är bärare av polysemi.

Resultatet kan bl.a. tänkas bli systematiska beskrivningar av hur särspråks ordförråd skiljer sig åt i polysemiskt avseende. Om man t.ex. jämför engelskans och svenskans lexikon kan man märka en tendens att engelskan oftare och allmännare transitiverar rörelseverb som beskriver kroppsrörelse. I termer av semantisk komponentanalys, kompletterad med semantisk valensteori, skulle man kunna uttrycka denna förändring så, att man byter ut (det underförstådda) objektet för kroppsrörelsen, nämligen den egna kroppen (som är identisk med subjektet och därför normalt inte uttrycks; ev. kan man använda reflexivpronomen), mot ett annat objekt. Detta objekt kan beskrivas som 'en kropp som inte är identisk med subjektet'. Det kan i andra fall röra sig om något som inte lever, men som i vissa avseenden liknar en kropp. Enligt Longman är det helt etablerat på engelska inta bara att säga *he walks*, utan också *he walks his dog* och t.o.m. *he walks the ladder* (när man förflyttar stegen på ett sätt så att det ser ut som om den går). På svenska är *han promenerar sin hund* väl marginellt möjligt, men knappast *?hon promenerar stegen*. Om vi ändå förstår det sistnämnda exemplet betraktas det nog som en frisk och tillfällig metafor. Det finns fler exempel på skillnader mellan svenskans och engelskans lexikon av samma karaktär som denna.

En annan mer övergripande skillnad mellan svenska och engelska tycks vara att svenskan har en större benägenhet att låta verbalsubstantiv syfta inte bare på en handling utan också på medel för den eller resultatet av den. Det svenska ordet *betalning* är polysemt mellan handling och medel. I engelskan används olika avledningsändelser för de två betydelseerna - *paying* respektive *payment*. Också i andra liknande fall kan samma förhållande mellan språken observeras. Det rör sig dock bara om tendenser och frekvensskillnader, inte om någon absolut omöjlighet för engelskan att använda verbalsubstantiv polysemt. *Drawing* betecknar exempelvis både handling och resultat, precis som svenskans *teckning*.

Om vi flyttar uppmärksamheten från tendenser som gäller hela semantiska fält till skillnader mellan ekvivalenter inom ett semantiskt felt, hamnar vi i förhållanden som tycks vara mer slumpmässiga. Eventuellt kan de ibland betraktas som kulturberoende.

Både engelska och svenska skiljer på *cap* / *mössa* respektive *hat* / *hatt*. Men i svenskan är det *hatt* som används om företeelser med liknande form och funktion, t.ex. svamphatt och hatten till en tub. I engelskan är det i stället *cap* som är polysemt på detta sätt. Om orsaken till denna skillnad är en slump, eller om det beror på engelska och svenska mössors och hattars utseende är omöjligt att säga. En användare som söker uppgifter om lexikaliska polysemier (i ett främmande språk) skulle i sådana fall kunna vara mer hjälpt av en ordbok som är sorterad i semantiska fält med anknuten nätverk av polysemier i stället för på det traditionella sättet med ekvivalenter till grundbetydelser som enda utgångspunkt.

I exemplen med transitivering av vissa rörelseverb såg vi hur svenskan var mindre polysem än engelskan. I svenskans lexikon blir det alltså en "lucka" för vissa betydelsemoment jämfört med engelskan. Det omvända förhållandet gäller förstås också ibland, t.ex. när svenska verbalsubstantiv är mer polysema än engelska. De ovannämnda luckorna lycktes delvis vara uttryck för en systematisk skillnad mellan språken. Andra luckor rör bara enstaka ord. Det finns också "delvisa" luckor, dvs. betydelser som endast delvis stämmer överens. Komponentanalysen passar bra också för att beskriva vari dessa delvisa likheter och skillnader består. Delvis överlappning är f.ö. något som också kännetecknar många grundläggande ekvivalentpar, särskilt när man kommer utanför de konkreta föremåls räjong.

I en första omgång av projektarbetet har vi koncentrerat oss på verbpolysemi. Vi har valt prototypverb och frekventa verb ur vart och ett av de semantiska fält som Viberg presenterat (bl.a. Viberg 1990). Den (delvis hierarkiska) analysmodell vi har utarbetat innehåller ett begränsat antal semantiska komponenter som fördelas på tre huvuddelar av verbens betydelse: allmän referens, specifik verbpreferens och semantisk valens (relationell betydelse). Inom var och en av dessa huvuddelar finns ett antal variabler som kan anta olika värden. (Inkodning sker i ett s.k. attribut-värde-system). Varje polysem betydelse ges värden för de olika variablerna i detta system. Dessa värden utgör viktiga komponenter i betydelsen. På så sätt kan man se dels vad som skiljer olika polysema betydelser åt, dels på vilka sätt två språks lexikon är lika och olika. Som komplement till modellen finns möjlighet att ge fria kommentarer, som bl.a. kan bestå i finare semantiska distinktioner, frekvensuppgifter, uppgifter om grammatiska företeelser m.m.

I. De allmänna komponenter som varit aktuella vid verbanalysen är följande:

POTENTIELL/REELL

Ett kriterium för att välja en av dessa betydelser är att den förstnämnda kan kombineras med *möjligt* i en sats, den andra med *faktiskt*.

GENERELL/AKTUELL

Om verbbetydelsen är generisk kan den parafreras med *brukar*, för aktuell betydelse är en sådan parafra inte möjlig.

KONKRET/ABSTRAKT

Konkret verbbetydelse antas föreligga om något av verbets valensled kan uppfattas med något av människans fem sinnen. Annars räknas betydelsen som abstrakt.

TIDSLIG/ICKE-TIDSLIG

Verb med reell och aktuell betydelse (de flesta verb) kan också ha tidslig betydelse. De kan då kombineras med tidsadverbial som anger tidsutsträckning eller tidpunkt. Icke-tidsliga verb kan inte kombineras med sådana adverbial.

DURATIV/MOMENTAN

Om verbbetydelse är durativa implicerar de tidsutsträckning och kan bl.a. kombineras med *länge* (Ekberg 1989:105). Momentana verbbetydelse kan kombineras med *plötsligt*.

Dessutom tillkommer komponenter från något av ett tjugotal semantiska fält, exempelvis LOKATIV, ÄGANDE, EXISTENS, VERBAL KOMMUNIKATION (Viberg). Kriterierna rör kombinerbarhet och parafrazyper. För att t.ex. lokativ betydelse ska sägas förekomma skall man kunna parafrasera en sats där den aktuella verbbetydelsen ingår med en sats som innehåller *där*, *därifrån* eller *dit* (POSITION resp. RIKTNING (KÄLLA/MÅL)). Ytterligare finindelningar i materialet kan göras efter önskemål.

II. Till de specifika verbvariablerna räknas endast skillnaden mellan status och process samt de olika aktionsarterna:

STATUS/PROCESS

Verbbetydelse som är statiska kan parafraseras med *ha* eller *vara*. De kan ej befrågas med *Vad hände?* (Platzack 1979:68). Processer kan däremot befrågas på det sättet.

Aktionsarter

ICKE-GRÄNSRELATERAD kan kombineras med tidsadverbial inledda av *i* (Andersson 1977:82).

GRÄNSRELATERAD kan kombineras med tidsadverbial inledda av *på* (Andersson 1977:82).

NEUTRAL (lexikalt obestämd mellan gränsrelaterad och icke-gränsrelaterad) kan kombineras med tidsadverbial inledda av både *i* och *på*.

ITERATIV (en kombination bestående av en icke-gränsrelaterad serie av gränsrelaterade skeenden) kan kompletteras med *regelbundet* (Platzack 1979:78f).

INGRESSIV/TERMINATIV

Gränsrelaterade (och neutrala) durativa verbbetydelse kan fokuseras på handlingens början och då parafraseras med *börja*, eller på handlingens slut och då kompletteras med *färdigt* (Norén 1990:76).

III. Valensleden och deras komponenter

Antalet valensled i modellen har hållits nere. Vi skiljer huvudsakligen mellan Agent, Upplivare och Objekt (två typer). Instrument anges också i vissa fall, även om det inte ingår i valensen.

Agenten betraktas som intentionell (vilket förutsätter bl.a. animathet). Kriteriet på INTENTION är att verbfrasen kan förekomma i imperativ (med agenten underförstådd) (Ekberg 988:77). Förekomsten av en agent säger alltså något om verbets betydelse. Icke-intentionella animata subjekt kallas Upplivare. Icke-intentionella, icke-animata subjekt kallas bara Subjekt.

Objekten indelas i sådana som påverkar aktionsarten (hos neutrala verb) och sådana som inte gör det. Till den förstnämnda kategorien räknas Lokation (rumsadverbial) och Tid (tidsadverbial) om de ingår i valensen och påverkar aktionsarten. De viktigaste betydelsekomponenterna hos aktionsartspåverkande objekt är AVGRÄNSANDE/OAVGRÄNSANDE. Avgränsande objekt kan kompletteras med *många* (jfr Platzack 1979:78f), oavgränsande med *mycket*. För övriga objekt finns en mängd komponenter som kan vara av intresse, t.ex. ANIMAT/INANIMAT, KONKRET/ABSTRAKT och AVGRÄNSAD/OAVGRÄNSAD. Dessa komponenter används vid behov för att skilja polysema betydelser eller två ekvivalenter åt. De kan också användas för att beskriva de aktionsartspåverkande objekten.

Instrument ges också komponenter efter behov.

Som exempel på hur modellen används beskrivs nedan analysen av ekvivalentparen *saw* - *såga* respektive *see* - *se*. För de polysema betydelsemomenten utöver utgångsbetydelsen anges endast vilka komponenter som avviker och vilken typ av avvikelse det är fråga om. Utgångsbetydelsen är i princip den betydelse som innehåller flest komponenter. De olika polysema betydelserna relateras till den betydelsen. Att de presenteras i en viss ordning är oviktigt; det viktiga är hur de förhåller sig till varandra. Beskrivningen är ett första försök; den kan säkert förbättras.

Svenska

1. *såga* Ex. *såga ved*

REELL

AKTUELL

KONKRET

TIDSLIG

DURATIV

FYSISK KONTAKT

(LOKATIV)

NEUTRAL

PROCESS

Agent

Objekt: föremål som ska delas

Engelska

1. *saw* Ex. *saw wood*

Samma komponenter som i svenska

/avskiljas

Instrument: såg e.d.

2. *såga* Ex. *de sågade sig elegant igenom försvaret.*

Omhierarkisering så att det lokativa elementet blir viktigare.

Tillägg (precisering) av komponenten LOKATIV med RIKTNING > FÖRFLYTTNING.

Byte av Objektinnehåll från föremål till den egna kroppen.

Förlust av Instrument.

2. ?

3. *såga* Ex. *såga på fiolen.*

Tillägg (precisering) av komponenten NEUTRAL med ICKE-GRÄNSRELATERAD.

Byte av Objektinnehåll till ett föremål som ej ska delas.

Tillägg (precisering) av Objektets funktion till OAVGRÄNS.

Byte av Instrument till något sågliknande i utseende (och ljud) men ej liknande i funktion.

3. *saw* Ex. *He was sawing away at his fiddle*

Samma komp. som i svenska.

4. (*Den här veden är lätt att såga*)

-

4. *saw* Ex *this wood saws easily*Byte från reell till potentiell betydelse
Därigenom icke-tidlig m.m.5. *såga* Ex. *juryn sågade henne /hennes bidrag*

Byte från konkret till abstrakt betydelse. Tillägg (precisering av) neutral aktionsart till GRÄNSRELATERAD, TERMINATIV.

Tillägg inom Objektet - det som avskiljs är icke önskvärt (och av speciell typ) samt AVGRÄNS.

Förlust av Instrument.

5.

-

jfr *spola*.

1. *se* 'uppfatta med synen'

Ex. *se en skymt av något.*

REELL

AKTUELL

KONKRET

TIDSLIG

PERCEPTION

DURATIV

ICKE-GRÄNSRELATERAD

PROCESS

Upplevare: (ICKE-INTENTION,
ANIMAT)

Objekt: KONKRET

Instrument: ögon

1. *see*

Ex. *see something in the fog.*

Samma komp. som i svenska

2. *se* 'titta'

Ex. *se (på) TV; se sig i spegeln*

Byte från Upplevare till Agent

(dvs. från ICKE-INTENTION
till INTENTION)

2. *see*

Ex. *For more information, see p. 5*
samma komp. som i svenska

3. *se* 'kunna se'

Ex. *hon ser inte så bra*

Byte från REELL till POTENTIELL

3. *see*

Ex. *he doesn't see very well*
samma komp. som i svenska

4. *se* 'uppleva, erfara'

Ex. *Huset har sett bättre dar*

Förlust av Objektsprecisering

Byte av Instrument från ögon

till alla sinnen (eller själen)

(I det svenska exemplet även
animering av subj.)

4. *see*

Ex. *we have seen some good*
times together

Samma komp. som i svenska

5. *se* 'möta, träffa'

Ex. *hon vill aldrig se honom mer*

Neutralisering så att Upplevaren

också är Agent. Byte från ICKE-

GRÄNSREL. till GRÄNSREL.

Byte från PERCEPTION till SO-
CIAL KONTAKT

5. *see*

Ex. *the doctor can't see you yet*

Samma komp. som i svenska

Tilllägg till Objekt: ANIMAT
Förlust av Instrument

6. se 'bjuda'

Ex. *se någon på en kopp kaffe*

Byte från PERCEPTION till
ÅGANDE. Tilllägg av LOKATION,
preciserad till RIKTNING mot
ett MÅL.

Byte från ICKE-GRÄNSREL.
till GRÄNSREL.

Byte från Upplevare till Agent
Förlust av Instrument.

6. ?

7. se 'tänka sig; föreställa sig'

Ex. *som jag ser saken*

Byte från KONKRET till ABSTRAKT.

BYTE från PERCEPTION till
KOGNITION.

Objektet byts ut och preciseras
till tankar o.d.

Byte av Instrument: hjärnan.

7. see

Ex. *the way I see it*

Samma komp. som i svenska

8. see 'inse; förstå'

Ex. *hon kunde inte se meningen
med det hela*

BYTE från KONKRET till AB-
STRAKT.

Byte från PERCEPTION till
KOGNITION.

Byte från ICKE-GRÄNSREL.
till GRÄNSREL.

Objektet byts ut och preciseras
till tankar o.d.

Byte av Instrument: hjärnan

8. see

Ex. *now I see what you mean*

Samma komp. som i svenska

9.

-

9. see 'follow'

Ex. *someone ought to see the
children safely home*

Byte ICKE-GRÄNSREL. till
NEUTRAL.

Byte från PERCEPTION till

LOKATION,

som preciseras genom tillägg av

RIKTNING mot ett MÅL.

Byte av Upplevare till Agent.

Byte av Objektinnehåll till ANIMAT.

Förlust av INSTRUMENT.

Genomgången av likheter och skillnader mellan de två ekvivalentparen ovan är inte gjord i projektets inkodningsformat utan på fri hand. Vid den riktiga inkodningen kopplas materialet till Språkdatas övriga material (via Ingres). Det kan sedan användas som hjälpmedel för att förbättra en- och, framgör allt- flerspråkiga ordböcker. Ett problem med sådana ordböcker för närvarande är att de i stor utsträckning överlåter åt läsaren att tolka polysema betydelser riktigt genom att de endast antyder sökriktningen. Angivelser som *äv. abstrakt* och *äv. bildligt/metaforiskt* är ganska lite informativa. Utgångsantagandet tycks vara att polysemirelationer, inklusive metaforik, är universella och därför begripliga för alla människor. Att detta inte är ett oproblemiskt antagande hoppas jag att ovanstående genomgång har visat. I förlängningen kan en noggrann och detaljerad genomgång av vilka olika betydelseförändringar som kan ligga bakom t.ex. en etikett som *metaforiskt* bidra också till att förenkla översättning, inklusive automatisk sådan.

Litteratur

Andersson, E. (1977) *Verbfrasens struktur i svenskan*. Åbo.

Ekberg, L. (1989) *Gå till anfall och falla i sömn*. Lund (Lundastudier i nord. språket. A:43).

Longman Dictionary of Contemporary English. (1987) Harlow, Essex.

Norén, K. (199). Svenska partikelverbs semantik. Inst. för svenska språket. Göteborg. (Stencil).

Platzack, C. (1979) *The semantic interpretation of aspect and aktionsarten*. Dordrecht.

Svensk ordbok (1986) Stockholm.

Viberg, Å. (1990) Svenskans lexikala profil. I: *Svenskans beskrivning* 17. Åbo akademi.

Datamatstøttet vending af ordbøger

Anders Drejer Nygaard

Denne artikel skildrer arbejdet med automatisk at producere en råudgave af en aktiv ordbog ud fra en tilsvarende passiv. En mærkning af den passive ordbogs indholdsstruktur efter den internationale standard SGML gør selve vendingen til en relativt simpel mekanisk proces. Et væsentligt delproblem er herefter at vælge hvilket opslagsord, udtrykkene i den aktive ordbog skal anbringes under, og en heuristik bliver beskrevet. Ideerne i denne artikel er blevet anvendt på Engelsk-færøsk ordbog, og Færøsk-engelsk ordbog er nu under produktion ud fra resultaterne.

Baggrund

Færøsk er, som bekendt, et lille sprog, så det store og bekostelige arbejde med at producere en ordbog fra grunden kan sjældent svare sig. Forlaget Stíðin besluttede derfor at købe manuskriptet til den røde engelsk-danske ordbog fra Gyldendal, og oversætte den danske del til færøsk, for herved at få en engelsk-færøsk ordbog. En tilsvarende fremgangsmåde lader sig ikke praktisere, når man vil producere en færøsk-engelsk ordbog. Imidlertid rådede man allerede over en maskinlæsbar udgave af en engelsk-færøsk ordbog, så hvorfor ikke prøve om det kunne lade sig gøre ad maskinel vej at 'vende' denne ordbog til den ønskede færøsk-engelske? Det var naturligvis fra starten klart, at der ikke ville kunne komme et færdigt produkt ud af anstrengelserne, men en stor del af indsamlingsarbejdet, og arbejdet med at finde ækvivalenter måtte kunne foretages på denne måde.

Fremgangsmåde

Centralt for arbejdet med at vende ordbogen er det begreb som Honselaar & Elstrodt [2] kalder en 'micro entry', men som jeg foretrækker at kalde en atomar artikel eller et atom. En atomar artikel indeholder præcis et par af (oversættelses)ækvivalente enkeltord eller udtryk, sammen med oplysning om emneområde, brugsoplysninger, eksempler, etc.

Ud fra denne synsvinkel er en ordbog en stor samling af atomare artikler, der er organiseret ved at man udnævner et af de to sprog til kildesprog (og dermed det andet til målsprog), sorterer indgangene efter kildesprogsord, og udfaktorerer information, der er fælles for flere atomare artikler. Denne information er så yderligere suppleret med

krydshenvisninger, kildesprogsforkortelser, begreber, der kun findes på kildesproget, etc.

Fremgangsmåden for at vende en ordbog er nu i princippet ganske enkel, og består af 6 trin:

- 1 Analyse
- 2 Udtynding
- 3 Opløsning
- 4 Vending
- 5 Sortering
- 6 Samling

Fase 1, Analyse: Ordbogen analyseres til struktureret form.

Dette er en velstuderet proces med en lang række kendte problemer (se fx Christensen [1] eller Jørgensen [3] for en nærmere beskrivelse). Resultatet af denne proces er en hierarkisk struktureret logisk opmærkning af hver artikel, således at der er klart afgrænset Hoved, OpslagsOrd, Krop, SubArt(ikel), Ækviv(alent), TydSpec(ifikation), etc.

Fase 2, Udtynding: Al supplerende information (henvisninger, parafraser, kildesprogs-specifikke forkortelser, etc.) bliver kasseret

En simpel proces, givet den logiske (modsat typografiske) opbygning af SGML-beskrivelsen.

Fase 3, Opløsning: Artiklerne opløses i atomer.

En simpel mekanisk proces, givet den hierarkiske opbygning af SGML-strukturen. Det eneste, der skal foretages er at kopiere hovedet og relevante TydSpecifikationer for hver Ækvivalent.

Fase 4, Vending: Alle atomer vendes (dvs. det tidligere målsprog udnævnes til nyt kildesprog).

Simpel, givet visse forudsætninger. I en ordbog forventer man at se ordklasseangivelser for opslagsordene, men under vending af Ensk-føroysk har der kun været adgang til ordklasseangivelser på de engelske ord. Dette er dog i praksis kun et lille problem, idet langt de fleste ækvivalentpar har samme ordklasse.

Fase 5, Sortering: Indgangene sorteres efter det ny kildesprog.

Skal for at være anvendelig også tage hensyn til alfabetiseringsreglerne for færøsk, samt kunne håndtere flere nøgler - fx har vi ønsket at den vendte ordbog havde alle udtryk og fraser samlet sidst i artiklerne, og anvendt dette som et af sorteringskriterierne.

Fase 6, Samling: Fælles information udfaktoreres.

En noget besværlig, men stadig mekanisk fjernelse af fælles information fra de atomare artikler. Resultatet af dette trin er en SGML-beskrivelse af den vendte ordbog, som kan redigeres videre.

I praksis er der dog adskillige komplikationer. Den principielle beskrivelse af en ordbog som en samling af atomare artikler er kun holdbar til en vis grænse - en ordbog er i reglen skrevet med henblik på fuld dækning af kildesproget, hvorimod dækningen af målsproget i høj grad vil være tilfældig, især når der findes flere synonyme eller nærsynonyme ord på målsproget. Dette, samt mangelen på begreber, forkortelser, etc., der er specifikke for målsproget, gør at resultatet (den vendte ordbog) i det mindste vil skulle suppleres med indholdet af en ordliste.

Primære ord i udtryk

Et centralt problem ved sorteringen af de atomare artikler er, hvordan man behandler kildesprogsudtryk. Leksikografens sædvanlige fremgangsmåde er at udnævne et af ordene i et udtryk til det primære, og anbringe oversættelsen sammen hermed. Denne udnævnelse vil imidlertid benytte sig af vage kriterier som 'det mest betydningsbærende ord' eller 'det mest karakteristiske ord' i udtrykket; noget end ikke menneskelige leksikografer kan blive enige om (som enhver der har prøvet at slå op i en ordbog vil vide!).

Nogle ting kan man dog sige om hvilket ord i et udtryk, der er det primære; det vil sjældent være en artikel, præposition, pronomen, hjælpeverbum eller konjunktion.

For en datamat er det således vanskeligt at finde det primære ord i et udtryk, men det er i nogen grad muligt at støtte sig til konteksten i den artikel, atomet stammer fra. Lad os tage et eksempel: I Ensk-Føroysk ser artiklen for substantivet 'keel' således ud:

I. keel *n* kjølur (eis *plfr*); kólpramur; on an even ~ á rættkjøl; (*flm*) rólīga, javnt og samt; lay (down) the ~ for strekkja kjølin til.

Her har vi 6 atomare artikler, der alle har 'keel' som det primære ord på engelsk:

keel - kjølur
 keel - kólpramur
 on an even keel - á rættkjøl
 on an even keel - rólīga
 on an even keel - javnt og samt
 lay (down) the keel for - strekkja kjølin til

Tre af de færøske oversættelser består kun af et ord, der naturligvis er det primære. Udtrykket 'á rættkjøl' består af to ord, hvor det ene (á) er en præposition, så her må 'rættkjøl' være det primære. I udtrykket 'javnt og samt' kan vi kun udelukke konjunktionen 'og', og får en uopløselig tvetydighed. I udtrykket 'strekka kjølin til' kan vi kun udelukke præpositionen 'til', men af konteksten fremgår det at 'kjølur' er en mulig oversættelse af det ord, 'keel', som forfatteren har valgt som det primære i det engelske udtryk, så derfor må den bøjede form 'kjølin' være et bedre valg end 'strekka' som det primære ord.

Da færøsk er et sprog, hvor ordene bøjes flittigt (bøjningssystemet minder en del om det tyske: der er fire kasus for substantiver; verberne bøjes i person), vil ordene i et udtryk i reglen optræde i en bøjet form, så for at kunne genkende oversættelser af det pri-

mære engelske ord i de færøske udtryk må man kunne genkende bøjede former. En egentlig morfologisk analyse af ordene i de færøske udtryk falder imidlertid uden for projektets rammer, så jeg har måttet nøjes med en heuristik.

Det viser sig at hvis man betragter ord, hvor de første tre/fire bogstaver er ens, får man tilfredsstillende resultater. Denne heuristik kan naturligvis ikke tage højde for alle bøjningsfænomener (vokalforskydninger er et oplagt eksempel), og den giver også en del 'falske' bøjningsformer (typisk vil sammensætninger blive klassificeret som bøjede former), men disse fejlkilder er forsvindende i sammenligning med problemerne i Fase 1, Analysen.

Algoritme til udvælgelse af primære ord

Svarende til observationen ovenfor om hvilke ord, der bliver valgt som primære ord, benytter algoritmen sig af en stopliste, der indeholder ord, der kun med ringe sandsynlighed kan være det primære ord i et udtryk.

Stoplisten indeholder alle præpositioner (á, av, ...), artikler (alla, alt, ...), konjunktioner (at, ...), almindelige hjælpeverber (fáa, fara, ...), pronominer (eg, okkum, ...), diverse almindelige småord (ikki, ið, ...), samt en række almindelige adjektiver (gomul, litil, ...). Ialt er der 180 ordformer på stoplisten.

Enhver ækvivalent inden for en artikel bliver behandlet efter tur, idet man forsøger at finde primære ord. Hvis ækvivalenten kun indeholder ét ord, eller hvis den kun indeholder ét ord, der ikke optræder på stop-listen er det primære ord *entydigt*. Hvis der ikke er et entydigt primært ord, tages i prioriteret rækkefølge enten:

- alle tidligere forekommende entydige ord i denne artikel,
- alle ord, der er bøjningsvarianter af tidligere forekommende entydige ord i denne artikel,
- alle ord, der ikke optræder på stoplisten, eller
- alle ord.

Bemærk at fremgangsmåden i praksis garanterer, at der faktisk kommer artikler i den vendte ordbog, der omhandler ord fra stoplisten - de skal bare forekomme som eneste ord i en oversættelse, hvilket de typisk gør i netop de relevante artikler i den originale ordbog.

Vending af Ensk-føroysk orðabók

Filosofien bag produktionen af den endelige skitse har været, at det skal være så let som muligt for redaktørerne at fuldføre processen. Da det er meget lettere at slette i et forlæg, end at tilføje manglende informationer, har vi valgt at lade tvivlstilfælde falde ud til fordel for overgenerering.

I Ensk-Føroysk orðabók forekommer der følgende oplysningstyper, som behandles som angivet:

- Opslagsord
Genanvendes som oversættelser
- Udtale
Slettes, idet udtaleangivelser for målsproget ikke ønskes
- Ordklasse (engelsk)
Genanvendes som færøsk ordklasse, i de tilfælde hvor ækvivalenten kun består af ét ord. Fejlprocenten ved at gøre dette er ganske lav (ca. 1-2), til gengæld for en stor besparelse i redaktionsarbejdet.
- Diskriminatorer
Genanvendes som diskriminatorer. Dette vil ikke altid være tilstrækkeligt (end ikke korrekt), men anvendeligt som udgangspunkt.
- Oversættelser
Genanvendes som opslagsord og udtryk.
- Parafraaser
Problematiske, idet de ikke altid er til at skelne fra egentlige oversættelser. I de tilfælde hvor en entydig identifikation som parafrase er mulig bliver oplysningen slettet, ellers bliver de behandlet som almindelige oversættelser.
- Eksempler (engelsk)
- Eksempler (færøsk)
- Forkortelser (engelsk)
Slettes, hvis der ikke er en færøsk ækvivalent
- Forkortelser (færøsk)
- Henvisninger
Slettes
- Rammebundne præpositioner
Overføres, idet de bliver vendt
- Brugsoplysninger
- Latinske navne
Bliver uddraget til en speciel fil, således at redaktøren kan behandle dem samlet.
- Typografisk hjælpeinformation: tegn, parenteser, nummereringer, etc.
Anvendes kun til at styre analysefasen

Den mængde af atomer, der bliver produceret ved analysen af en ordbog som Ensk-førøysk, udgør en temmelig stor 'seddelsamling', som i fase 6 bliver organiseret til noget, der ligner en ordbog ved at bortfaktorerer fælles information. Denne 'ordbog' skal, ud over indholdet i den endelige ordbog, også indeholde angivelser af, hvornår der har været flertydighed ved bestemmelsen af det primære ord - der er jo ingen grund til at have den samme oversættelse stående flere gange i samme ordbog. Dette har vi valgt at gøre ved at understrege ord i færøske udtryk, som er alternative placeringer for det pågældende udtryk.

Endvidere har vi ment, at det kan være praktisk for redaktøren at kunne se, hvorfra en given oversættelse stammer, og vi har derfor indført en '*' ved det engelske opslagsord i de tilfælde, hvor der kan opstå tvivl.

Konklusion

De 47.555 artikler i Ensk-føroysk orðabók er blevet vendt til 64.610 artikelskitser (her er bøjede former talt flere gange).

Der har været en enorm besparelse med hensyn til indsamling af oplysninger.

Der er en meget lille besparelse på redigeringen af de store artikler (se fx 'royna', figur 1), men mange af de helt små artikler er faktisk færdigredigerede.

Endelig har hele processen afsløret en del mangler, inkonsekvenser og egentlige fejl i forlægget, Ensk-føroysk orðabók - især alfabetiseringen af den vendte ordbog har vist sig at være nådesløs til at afsløre stavfejl i forlægget.

Man må således sige, at hele processen har medført en stor arbejdsbesparelse, samt givet et nyt og værdifuldt indblik i en eksisterende ordbog, altsammen for en relativt beskedne indsats.

Litteratur

1. Ole Norling Christensen: *Struktureret redigering af ordbøger*, i Nordiske studier i leksikografi, Nordisk forening for leksikografi, Oslo 1992.
2. Wim Honselaar og Marijke Elstrodt: *The electronic conversion of a dictionary: from Dutch-Russian to Russian-Dutch*, i proceedings from EURALEX '92, Department of Translation Studies, University of Tampere, Tampere, Finland 1992.
3. Claus Bo Jørgensen: *Parsing af ikke-strukturerede ordbogsdata*, i proceedings fra Konference om Leksikografi i Norden, 1993

royn: kom og royn come and have a *try; royn aftur have another *try; royn eina fer afturat have another *try; royn ikki at lumpa meg don't *try anything on with me.

royna hj *out to (td make money); #. n put to *trial; T have a *smack at; T try-out; (ein el eitt) put to the *proof; (ein) put to the *test; (seg við) have a *go at; #. s *check up on; *go over (td let us *go over the last act); *try out; assay; attempt; endeavour; essay; experience; experiment; find; fish; offer (td he -ed to strike me); prove; prove (td sannleikaviri d); see; seek; taste; test; try (td his patience was tried); (glm) prove (td his worth); (um d-pd, flm) sound; (um skotvápn) range; (vi- snøri el trá-u) angle

· fara at royna eitt T do it on *spec (st speculation) (td I don't know whether he is there, but I'll go there on *spec); eg fari at royna at náa tokinum I'll have a *shot for the train; lat hann sleppa at royna seg (el at royna flogi-) give him a *trial; royna aftur repeat; royna at fáa *go for; royna at *study to; royna eitt have (el make) a *stab at sth; have a *whack at; royna hann put him through his *facings; royna seg við have (el take) a *fling at; royna seg ímóti (el í) T have a *bash at; royna seg ímóti *measure one's strength with; try *conclusions with; try a *fall with; royna seg try one's *luck; (flm um persón) expand; (ímóti) match (td *match your strength against his); royna at fáa at vita hvat hann hevur í kvinnal (num) (flm, eis) *sound him out; royna at fáa *try for; royna at vinna sær tí *play for time; royna at S have a *shot at sth; royna seg við try one's *hand at; ta- at royna eitt T give it a *whirl.

roynandi / tentative.

roynast s *prove oneself (to be) (td he -d himself to be a true friend); deliver; make; prove (td the story -d false); S *deliver the goods; (vi- l-singaror-i, væl el illa) run; (væl el illa) *turn out; (væl) *prove itself (td this method has -d itself); (væl) *prove oneself (td this method has -d itself)

· eg haldí meg vita hvussu hann fer at roynast I have got him *taped; hann fer at roynast he *improves on acquaintance; nógv bendit á at ta- fer at roynast væl show (great) *promise; roynast sannur *prove true; roynast sum ma-ur *play the man; ta- at roynast verri enn evnini eru til underachievement.

Blant fire millioner sedler

En situasjonsrapport fra Dokumentasjonsprosjektet

Christian-Emil Ore

Det store arbeidet som er lagt ned i oppbygningen og vedlikeholdet av seddelsamlinger vitner om hvilket viktig hjelpemiddel ordsedlene har vært i ordboksarbeid. I de senere år har datateknikken muliggjort en mye mer effektiv oppbygning av den informasjonen som et seddelarkiv representerer. Elektronisk lesning av tekst (OCR), konkordansprogrammer og hjelpemidler for (halv-) automatisk markering av grammatisk informasjon på ord i løpende tekst kan nevnes [se Atkins 1992]. Tilstedeværelsen av denne effektive teknikken gjør det fristende å konsentrere seg om kilder som for en relativt billig penge kan gjøres elektronisk tilgjengelig eller som allerede er på elektronisk form. En slik dreining kan lett medføre at verdifull informasjon i de manuelle arkivene blir neglisjert fordi den er tungt tilgjengelig.

Dette problemet er selvfølgelig ikke bare begrenset til leksikografien, men finnes innen en rekke humanistiske fag. Dokumentasjonsprosjektet er et fellesprosjekt ved de fire universitetene i Norge som skal bøte på noe av problemet. Det går i korthet ut på å omforme papirarkivene ved en rekke samlingsavdelinger (muséer o.l.) ved de humanistiske fakultetene til elektroniske form. Prosjektet vil gå frem til 1998, og målet er å bygge opp "Universitetenes databaser for språk og kultur". Basene vil i første omgang omfatte de 12-14 millioner arkivkort og dokumenter ved samlingsavdelingene, men vil senere omfatte store bildebaser samt elektroniske kart for å lette koblingen av de forskjellige databasene. Prosjektet vil forhåpentligvis muliggjøre en helt ny krysskobling av data ved de ulike samlingene.

Leksikografidelen av Dokumentasjonsprosjektet

Innen leksikografi finnes det i Dokumentasjonsprosjektet tre delprosjekter, et for hvert av feltene bokmål, gammelnorsk og nynorsk. Dette avspeiler den tilsvarende tredelingen av avdeling for leksikografi. I alle de tre underavdelingene er det store ordseddel-samlinger, til sammen syv millioner sedler. Selv om den opprinnelige målsetningen for hele Dokumentasjonsprosjektet var "å gjøre seddelarkivene tilgjengelig på elektronisk form" er det ikke tale om å behandle mer enn fire millioner av disse sedlene. Dette skyldes at det finnes store mengder sedler som bare har den samme informasjonen som man vil finne i en vanlig KWIC-konkordans. Det ville være uansvarlig bruk av ressurser å skrive av disse sedlene. I stedet vil det bli brukt optiske lesere for å gjøre de opprinnelige

tekstene elektronisk tilgjengelige. Disse tekstene vil bli tagget i henhold til den SGML-baserte "Text Encoding Initiative" (Goldfarb 1991, Sperberg-McQueen and Bernard 1990)

Konverteringsarbeidet er nå i full gang for alle de tre avdelingene. Til sammen er det engasjert ni assistenter og omlag 100 ufaglærte innskrivere på hel- eller deltid. De sistnevnte er alle arbeidsledige som i enten deltar "Arbeid for trygd" tiltak eller er tilknyttet en av våre spesiallagde studie- og registreringsentraler i Nord-Norge.

Det aller meste av Bokmålsavdelingens sedler er av den nevnte KWIC-konkordanstypen, riktignok med ordklassemarkering. I tillegg til disse har avdelingen et seddelarkiv med opplysninger om omlag 350 000 ord i norsk sakprosa. Dette ekserptarkivet er nå skrevet inn og skal slås sammen med et tilsvarende, allerede elektronisk tilgjengelig arkiv påbegynt i 1968 over nyere norske ord i aviser og tidsskrift. Det arbeides også med å gjøre et utvalg verker av norske forfattere elektronisk tilgjengelig.

Gammelnorskavdelingens sedler er også av konkordanstypen, men skiller seg fra de litterære sedlene for moderne norsk i minst to henseender. Teksten sedlene er basert på, er en korrigert utgave av de tilgjengelige trykte tekstutgavene av manuskriptene. Det er derfor ikke mulig å laste inn de trykte tekstene. Det kompliserer oppgaven. På den annen side inneholder sedlene både linje og sidehenvisninger til den trykte utgaven. Det er derfor mulig å legge sedlene inn i et databasesystem på en slik måte at den originale teksten kun blir skrevet en gang (se også Ore 1991). Idag er de 65 000 sedlene basert på "Thomas saga erkebiskups" skrevet inn i dette systemet. Vi har dermed en løpende tekst på omkring 65 000 ord der hvert ord er påført grammatiske opplysninger og normalisert grunnform. Innskrivningsarbeidet ble utført av fire ufaglærte i løpet av et halvt år. Dette er en beskjeden innsats når man tenker på hvilket arbeid som ligger bak det å lage ord-sedlene. Når hele Gammelnorskavdelingens seddelsamling på i alt 700 000 sedlene er skrevet inn, vil vi ha en interessant base over grunnord og deres realisasjon i en rekke tekster. I tillegg til seddelarkivet arbeides det for tiden med å legge inn det samlede gammelnorske korpuset.

Det nynorske seddelarkivet representerer en blanding av det meste. Her finnes alt fra litterære belegg til sedler som kun gir opplysninger om at ord er hørt et sted i Norge. Arkivet er bygd opp over flere generasjoner og mange hundre informanter har bidratt. Å forsøke å laste alt dette inn i et formelt skjema kan synes vel dristig. Til dette er det å si at ikke alle sedlene vil bli skrevet inn. Vi frasorterer de fleste litterære sedlene og også alle sedler basert på dialektordlister. Disse vil enten erstattes av hele, elektroniske tekster eller vil bli skrevet inn under ett.

Innlegging av materialet

For gammelnorsksedlene har vi laget et innskrivningskjema som har vist seg meget effektivt. Ved de andre avdelingene har vi valgt å skrive inn hver enkel seddel som en løpende tekst med koder. Bokmålsavdelingen hadde allerede et ferdig kodeoppsett utviklet på 70-tallet. Så her var oppgaven også ganske enkel. For for den nynorske seddel-

samlingen eksisterte hverken en modell eller et kodeskjema. Da formaliseringsarbeidet startet våren 1992 hadde avdelingen en del erfaring i å kode ordboksmanuskripter. I samarbeid med Bokmålsavdelingen var det blant annet utviklet et feltskjema og et enkelt registreringsprogram (Felted) i forbindelse med utgivelsen av håndordbøkene "Nynorskordboka" og "Bokmålsordboka". I dette systemet finnes det lagret omkring 30 000 sedler med opplysninger om ekserpter fra sakprosa. Men materialet i det store arkivet viste seg å være mer sammensatt og uensartet enn de enkle formaliseringssystemene kunne takle. Formaliseringsarbeidet har strukket seg over vel ett år og det har vært mye prøving og feiling underveis. Men assistentgruppen på nynorskavdelingen har gjort et kjempearbeid (se Hagen og Ragnsæter 1993). Vi har nå både et SGML-taggesystem for arkivet og en formell datamodell. Det er også utarbeidet et hundresiders instruksjonshefte for koding av sedlene (Hagen, Haukaas og Ragnsæter 1993).

Datamodellen

Det første skrittet mot en database er å foreta en grundig analyse av de dataene som skal legges inn. Analysen bør resultere i en datamodell beskrevet i et formelt datamodelleringsverktøy. Dette kan riktignok gjøre jobben vanskeligere idet en formell beskrivelse må være pinlig korrekt. Men det gir en større sikkerhet mot inkonsistenser enn bruk av et uformelt feltskjema. Vi bruker datamodelleringsverktøyet NIAM (Nijssen 1988), men det finnes en lang rekke metoder (se Hjort & al 87).

I figuren er det tegnet et forenklet bilde av den formelle datamodellen for nynorsk-materialet. Ovalene representerer hovedkategoriene. Linjene indikerer forholdet mellom dem. "Buntene" betyr mange-til-mange relasjoner. Et oppslagsord kan være forbundet med mange ordformer, og en ordform kan stå i forbindelse med flere oppslagsord (f.eks. andre hovedformer eller sidestilte former). En enkelt linje forteller om en en-til-en forbindelse, mens de sprikende linjene angir en-til-mange forbindelser. Boksene er brukt for å antyde hva slags tilleggsinformasjon som er lagret om elementene i de enkelte kategoriene.

Den venstre delen av diagrammet viser sammenhengen mellom data i et tagget tekst-korpus der ordene er kodet med normert form og morfologiske opplysninger. Den nederste ovalen i diagrammet representerer samlingen av alle tekstene, mens den nest øverste ovalen representerer de enkelte ordformene slik man finner dem i tekstene. Forbindelsen mellom disse er gitt ved ordformens aktuelle plassering i en tekst. Den øverste ovalen er de normerte oppslagsordene slik man vil finne dem i en ordbok. I verdenen av ordsedler vil forbindelsen mellom denne og den nest øverste ovalen svare til forbindelsen mellom det normerte ordet øverst på ordseddelen og det understrekte ordet i seddelteksten. Sann sett beskriver den venstre delen av diagrammet hvordan opplysningene på sedlene i gammelnorsk- eller bokmålsarkivet henger sammen.

Nynorsksedlene er derimot mer komplekse. De kan ha illustrasjoner, det kan stå ekstra kommentarer osv. Dette er illustrert i diagrammet ved at vi har en egen kategori "seddel". Nynorskarkivet skal dokumentere både talemål og skriftspråk. Dette betyr at

det finnes mange sedler som viser at et ord er brukt et sted i Norge. På en slik seddel vil det normalt være et eller flere sitater som viser ordet i bruk, samt en unormert grunnform av det. Denne målføre-grunnformen vil skille seg fra både den normerte oppslagsordet og fra den aktuelle (bøyde) ordformen ieksempellet. Vi har derfor introdusert en tredje ordkategori kalt "målføre-grunnformer". Til denne knyttes opplysninger om sted, hva ordet betyr, bøyning

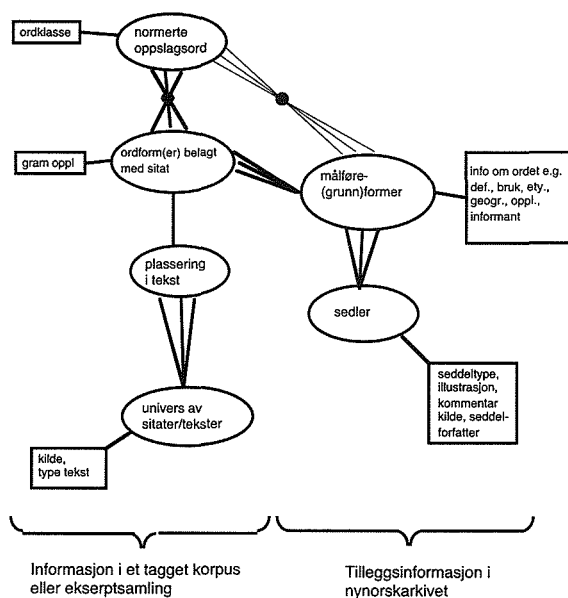


Fig. 1 En forenklet skjematisk fremstilling av informasjonen i seddelbasene

og liknende. Det hender også at det ikke står noe sitat på seddelen. Det er derfor en direkte forbindelse mellom målføre-grunnformen og den normerte utgaven av ordet. Kategorien "målføre-grunnformer" inneholder også sammensetninger og ordsamband. I tillegg finnes det sedler der målføre-grunnformen ikke er ført opp, men hvor det står opplysninger om bruk. Slike sedler blir kodet som om grunnformen eksisterte, men feltet vil bli stående tomt i databasen.

Seddeltvirkeligheten er adskillig mer kompleks enn det denne korte gjennomgangen kan gi inntrykk av. Men figuren gir hovedtrekkene i databasene som bygges opp på grunnlag av de eksisterende seddelarkivene (også bokmål- og gammelnorsk materialet). Diagrammet kan også illustrere at et (morfologisk tagget) tekstkorpus kan settes mer eller mindre rett inn i systemet. For folk som ikke er vant til å tenke i bokser og piler er dette kanskje ikke selvsynlig. Tekstene i et slikt korpuset havner under kategorien "Univers av sitat/tekster", mens de enkelte ordformene med grammatiske opplysninger går inn i ovalen "ordform(er) belagt med sitat". Elektronisk tilgjengelige ordbøker vil også kunne føyes til ved at oppslagsordene grupperes under "normerte oppslagsord".

Ordartiklene vil da bli informasjonspakker forbundet med disse. Slik kan man ad libitum fortsette å bygge opp en kompleks dataverden.

Videre arbeid

Leksikografidelen av Dokumentasjonsprosjektet har som hovedmål å lage elektroniske hjelpemidler som kan erstatte avdelingenes bruk av sedler. Idag er konverteringen av de eksisterende seddelsamlingene godt igang, mens oppbygningen av elektroniske tekst-samlinger er i startfasen. Slik sett går prosjektet etter planen. Det mangler imidlertid noe "å vise frem". Det meste av vår tid har vært brukt til å bygge opp innskrivningskapasiteten og metoder for formalisering av materialet. I de neste to årene vil vi arbeide mer med verktøyer og systemer for å bruke dataene i det praktiske, leksikografiske liv. Vi håper å kunne gi en positiv rapport om dette på konferansen i Reykjavik i 1995.

Litteratur

- Atkins, S. *The Hector Project*, i "Proceedings of Complex '92, Budapest 1992
- Goldfarb, C. *The SGML Handbook*, Oxford University press 1991
- Hagen, K., Haukaas, J., Ragnsæter, O. *Instruksjonshefte for innskriving og tagging av ord-setlar*, Oslo 1993
- Hagen, K., Ragnsæter, O. *Nynorskdelen av dokumentasjonsprosjektet* Språklig samling 2 Oslo 1993
- Hjort 87 The Danlex -Group, *Descriptive Tools for Electronic Processing of Dictionary Data* Lexicographica Series Maior 20 Max Niemeyer Verlag, Tübingen 1987
- Ore, C.-E. *Dokumentasjonsprosjektet ved Det historisk-filosofiske fakultet, Universitet i Oslo*, i Nordiske studier i leksikografi, konferanserapport, Oslo 1991
- Sperberg-McQueen, C.M., Bernard, L. (eds) *Guidelines for the Encoding and Interchange of Machine-Readable Texts (TEI P1)*, Chicago and Oxford Noveber 1990

Ordbog og encyklopædi

Karen Margrethe Pedersen

Ordbøger er opslagsværker om sprog, mens encyklopædier er opslagsværker om "verden". Men ordbøger indeholder ofte encyklopædiske oplysninger, og encyklopædier sproglige oplysninger. I det følgende skal jeg præsentere en ordbog, der medtager mange encyklopædiske oplysninger, og jeg skal vise, hvorledes disse oplysninger er indpasset i ordbogens struktur.

Ømålsordbogen

Ømålsordbogen beskriver som navnet siger dialekterne på Øerne. Dialektforskerne deler Danmark i 3 hovedområder: Jylland, Øerne og Bornholm. Dialekterne i Jylland bliver beskrevet i Jysk Ordbog, der redigeres ved Jysk Institut, Århus Universitet. Dialekterne på Bornholm er beskrevet i en ældre Bornholmsk Ordbog, og Øernes dialekter bliver beskrevet i Ømålsordbogen, der redigeres ved Institut for dansk dialektforskning, Københavns Universitet.

Jysk Ordbog og Ømålsordbogen hører til de såkaldte "store nationale ordbøger" i stil med Ordbog over det Danske Sprog. ODS beskriver rigsmålet fra år 1700 og frem til omkring 1950, og dialektordbøgerne beskriver dialekterne i samme tidsrum, men med hovedvægt på tiden 1850-1920. Dengang levede over halvdelen af befolkningen på landet og de fleste talte dialekt. Nu er det kun et mindretal, der taler dialekt; dialekterne er ved at blive historiske sprog, og dialektordbøgerne er historiske ordbøger.

I modsætning til ODS og andre rigsmålsordbøger bygger dialektordbøgerne helt overvejende på mundtlige kilder. Materialet er indsamlet ved optegnelse hos de mennesker, der talte dialekt. Dialektforskerne har siddet ude hos folk i bøndergårde og fiskerhuse og landarbejderhuse, med spørgelister og papir og blyant, eller i nyere tid med båndoptagere. Ordbøgernes materiale er således et unikt materiale; det indeholder et væld af oplysninger, som man ikke finder andre steder, ikke kun om sproget, men også om den verden, som sproget er knyttet til. Og denne verden er ikke længere almindelig kendt.

Ord og sag

Ømålsordbogen er præget af to vigtige retninger inden for sprogforskningen: strukturalismen, der blomstrede i Danmark i 1940'erne og 50'erne i kredsen omkring Louis

Hjelmslev, og Wörter-und-Sachen-forskningen, der blomstrede i Tyskland i 1920erne og 30erne. Strukturalismen har præget beskrivelsen af ordenes udtryksside og dermed udtaleafsnittene, Wörter-und-Sachen-forskningen har præget beskrivelsen af indholdssiden og dermed betydningsafsnittene. De to retninger har haft en afgørende indflydelse både på indsamlingsmetoderne, redaktionsreglerne og det færdige resultat.

Hovedtanken i ord-og-sag-forskningen er, at man ikke bør studere ord for sig og sag for sig, men at ord og sag hører tæt sammen og bør beskrives under ét. Det er en indlysende tanke for en redaktør af en énsproget ordbog: Ordene henviser til ting og begreber og handlinger, og det er ordbogens opgave at beskrive henvisningen. Det sker i betydningsangivelserne. Og ting og begreber og handlinger har "navne", og det er ordbogens opgave at registrere disse navne.

Vi kan ikke redigere et ord uden sag. Vi kan fx ikke redigere ordet *drivtov*, hvis vi ikke har optegnet, hvad et drivtov er. Vi kan naturligvis forsøge at slå ordet op i andre ordbøger for at finde en forklaring, men for det første er det betænkeligt, hvis en dialektordbog skriver af efter rigsmålsordbøger, og for det andet finder vi ikke ordet i nogen anden ordbog. Hvis vi kan se på optegnelsen, at ordet har med fiskeri at gøre, kan vi forsøge at slå op i forskellig faglitteratur om fiskeri. Men fagbøgerne har sjældent sans for terminologi. De beskriver tingene, men ikke deres navne, og den terminologi, der anvendes, er tit fagfolkernes egen fagterminologi, ikke den folkelige. Etnologer taler fx om *faststående garn* modsat *drivgarn*, mens fiskerne selv kalder de faststående garn for *sættegarn*. Vi kan finde mange billeder af tove i fagbøgerne, men vi får ikke at vide, hvad for ét af dem der hedder et *drivtov*. Den eneste løsning på problemet er at få fat i en gammel fisker og spørge ham, hvad ordet betyder.

Vi kan heller ikke redigere en sag uden ord. Hvis ingen nogen sinde har spurgt drivgarnsfiskerne om, hvad de kalder det tov, som de binder garnlænken fast til båden med, så vil ordet *drivtov* ikke komme med i ordbogen. Redaktøren vil ikke opdage, at ordet mangler i seddelkassen, og ordet vil være tabt for eftertiden.

Der hviler således et stort ansvar på optegneren. Den, der optegner fiskeri skal være godt inde i emnet. Han skal tage udgangspunkt i den saglige verden og minutiøst gennemgå alle redskaber og alle fiskeprocesser med den enkelte meddeler for at få hele ordforrådet med. Og han skal beskrive hvert ord tilstrækkeligt præcist og detaljeret til at ordbogsredaktøren senere kan definere det præcist. Og så er det ikke kun fiskeri, der skal beskrives på denne måde. I teorien er det hele den dialekttalendes verden.

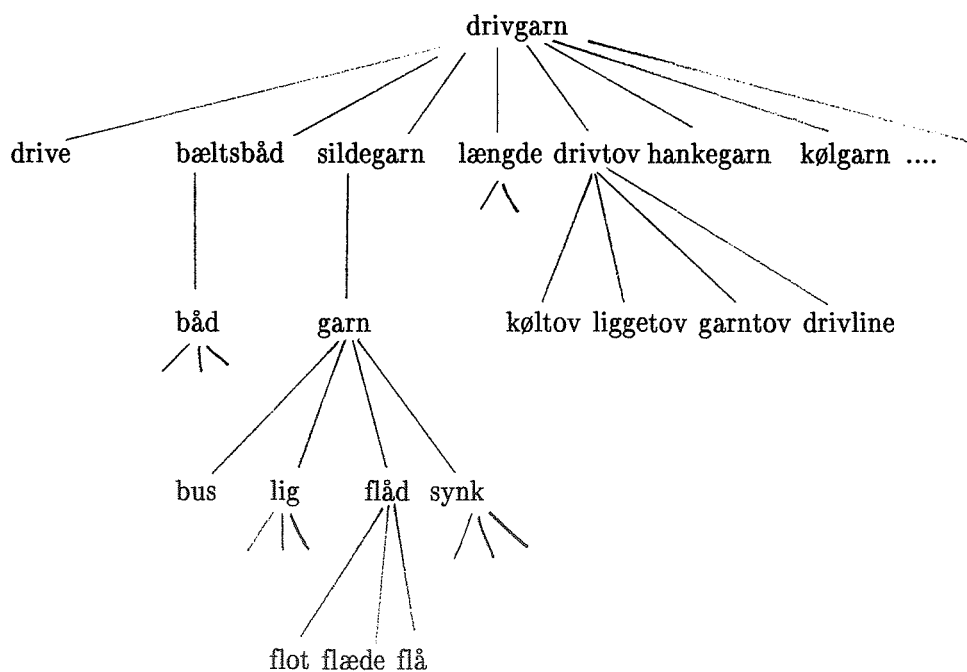
For at leve op til disse krav delte man verden imellem sig. Redaktørerne specialiserede sig i hver deres emner. En af redaktørerne specialiserede sig fx i fiskeri, han læste faglitteratur om fiskeri, og han tog rundt til fiskerlejerne og optegnede hos de gamle fiskere. Han specialiserede sig også i emner som pløjning, høst og tærskning. En anden redaktør tog sig af bagning og slagting og børns lege, og en tredje af gårdens bygninger og inventar osv. (Side 5 i Tillægsbindet til ØMO kan man se en liste over alle de "handlings-emner" og "begrebsområder", der blev fordelt mellem redaktørerne).

Leksikonafsnit

Ord-og-sag-forskningen bevirkede, at man foretog meget detaljerede optegnelser med udgangspunkt i den saglige verden. Men man gik videre end det. Man besluttede, at ordbogen skulle indeholde sammenhængende saglige beskrivelser af de vigtigste redskaber og de vigtigste arbejdsgange. Herved bliver ordbogen ikke kun en ordbog, men også et leksikon over den materielle side af bonde- og fiskerkulturen. Ordbogen overskrider grænsen mellem ordbogsgenren og leksikongenren.

Som eksempel på et leksikon-afsnit har jeg valgt artiklen *drivgarn* (se bilag 1). Artiklen indledes med en definition: "fiskegarn til at *drive* med (efter sild) (mods *sættegarn*)", en geografisk udbredelse: Ø(alm), og et par citater med lokalisering. Så langt ligner artiklen en almindelig ordbogsartikel. Men derefter følger der et leksikonagtigt afsnit om drivgarnsfiskeri: "drivgarnsfiskeri foregik om efteråret ..".

De dialektale termer i afsnittet er kursiverede, og kursivering er lig med henvisning. Under de kursiverede ord henvises der dels tilbage til *drivgarn*, dels frem til andre ord, hvorfra der igen henvises både frem og tilbage. Henvisningerne indgår således i et større eller mindre netværk.



Der henvises blandt andet:

- til synonymer (fra *drivtov* til *køltov*, *liggetov* og *garntov* og tilbage fra disse til *drivtov*) (*drivtov* er valgt som hovedsynonym i synonymkredsen)
- til antonymer (fra *drivgarn* til *sættegarn* og omvendt)

- fra underbegreb til overbegreb (fra *bæltsbåd* til *båd*) og omvendt (fra *båd* til *bæltsbåd*)
- fra en genstand til dens dele (fra *garn* til *bus*, *lig*, *flåd* og *synk*) og omvendt (fra *bus* osv. til *garn*)
- fra en handling til dens dele (fx i artiklen *bage*: fra *bage* til *ælte*, *slå* (*brød*) *op*, *smøre*, *gistne* m.fl.) og omvendt (fra *ælte* osv. til *bage*)
- fra handlinger til de genstande, der indgår i dem (fra *drive* til *drivgarn*) og omvendt fra genstande til de handlinger, de indgår i (fra *drivgarn* til *drive*, *give* (*garn*) *ud*, *sejle* (*garn*) *ud*)
- fra handlinger til de personer, der udfører dem (fx i artiklen *bryde*: fra *bryde* (*hør*) til *brydekone* og *brydemand*) og omvendt fra personer til de handlinger de udfører (fra *brydekone* og *brydemand* til *bryde*)

Hele netværket af krydshenvisninger gør, at ordbogen har en saglig indgang til sproget. Normalt har ordbøger kun en sproglig indgang: man kan slå op under ordet *drivtov* og se, hvad det betyder. Men i ØMO kan man også "slå op under sagen" og se, hvad den hedder. Fx kan man finde ud af hvad de forskellige drivgarn og tove hedder ved at slå op under *drivgarn*. Med en smule fantasi kan man i mange tilfælde gå ikke blot fra ord til sag, men også fra sag til ord.

Forskellen på ordbogens leksikonafsnit og almindelige leksikonafsnit er, at ordbogen anvender den folkelige terminologi. Man får hele tiden at vide, hvad tingene hedder, og det vil sige, hvad de hedder på Øernes dialekter.

Leksikonafsnit findes både under substantiver og verber. Under *bi* er der en beskrivelse af bier og biavl med tilhørende terminologi, under *bord* et afsnit om stuebordenes indretning og delenes navne. Under verbet *binde* har man ikke mindre end 3 saglige afsnit: et om at binde neg, et om at binde fiskegarn og et om at binde i betydningen strikke.

Det er den materielle folkekultur, der er i fokus i de saglige afsnit, dvs redskaber og arbejdsgange. Den åndelige folkekultur i form af folketro, sagn og eventyr hører under folkemindeforskerne, men ordbogen har dog en del stof vedrørende folketro.

Saglige citater

Saglige oplysninger findes ikke kun i særlige leksikonafsnit, men også i citater og definitioner. I artiklen *bleg* (bilag 2) ser man et eksempel på saglige citater.

Der er 2 typer af citater i ordbogen. Den ene type giver saglige oplysninger: "det lærer der *lå til bleg* blev vandet flere gange i døgnet", den anden type gengiver dagligdagens replikker: *I dag skal vi have tøjet på bleg*. Replik-citaterne er i kursiv, mens de sagligt beskrivende citater er i ordinær i anførselstegn. (Det sidste, saglige citat i artiklen *bleg* er en undtagelse: det er i kursiv, fordi det er i lydskrift, og lydskriften er født kursiv).

Replikken *i dag skal vi have tøjet på bleg* er konstrueret til lejligheden her, der er ingen replik-citater under *bleg*. De dagligdags replikker er i det hele taget relativt sjældne under substantiver og mere almindelige under verber og adjektiver. Substantiverne er sagligt mere interessante og fraseologisk mere uinteressante, sagt meget groft. Men fraseo-

logien er dog markeret i artiklen *bleg*, dels i den spatierede forbindelse *l i g g e, v æ r e, h æ n g e p å b l e g*, dels i de kursiverede dele af citaterne.

Den saglige tilgang til ordene kan undertiden føre til rent saglige citater. Det kan man se eksempler på i artiklen *garn* (bilag 3), i afsnittet om folketro: "når man ikke havde held med garnene, skyldtes det, at de var forheksede..". Citaterne her er atypiske for ordbogen: de bidrager ikke til at belyse opslagsordets betydning eller fraseologi, men falder helt uden for ordbogens genre. De burde måske have været udeladt.

Saglige definitioner

I artiklen *blusselygte* (bilag 2) kan man se et eksempel på, at saglige oplysninger gives i definitionerne. En blusselygte er en 'lygte til at blusse med'; det kan man næsten skrive uden at kende genstanden, men hvis definitionen skal være tilstrækkelig præcis, må man angive, hvad der menes med *blusse*. Det er gjort ved en henvisning til artiklen *blusse* (bet. 2.5, om at blusse ål). Så langt har vi at gøre med en almindelig ordbogsdefinition, men derefter kommer der en saglig, leksikonagtig tilføjelse: "først tran- ell tællelygte, senere petroleum- ell karbidlygte".

Den leksikonagtige del af definitionen er sat i parentes. Grænsen mellem ordbog og leksikon er markeret klart. Men grænsen er ikke altid så klar eller skarp. Det kan i nogle tilfælde bero på et skøn, om et givet træk eller karakteristikon ved det, der skal defineres, opfattes som et essentielt, definerende træk, eller om det opfattes som et accidentielt, ledsagende træk.

De saglige tilføjelser til definitionerne svarer på en måde til de saglige leksikonafsnit, men er mindre omfattende og indeholder ingen eller kun lidt terminologi. Og omvendt udgør leksikonafsnitene saglige tilføjelser til definitionerne, idet de beskriver de ledsagende træk. Som eksempel kan man tage artiklen *bage* (i ØMO bind 1): At *bage* betyder ifølge ØMO at 'fremstille brød eller kage ved stærk opvarmning i en ovn'. Det betyder det også på nudansk rigsmål, og i den forstand kan man sige, at sætningen *jeg bagte i går* betød det samme, når den blev sagt af min oldemoder på Sydfyn, som når jeg siger den. Men det arbejde, min oldemoder udførte, når hun bagte i den store murede bageovn, var meget anderledes end det arbejde, jeg udfører, når jeg bager. De ledsagende træk ved bagningen var forskellige. I den forstand betød *jeg bagte i går* noget andet for min oldemoder end for mig. Og det fremgår af ordbogens leksikonafsnit under *bage*, hvor den gamle bageproces beskrives. (Afsnittet fylder 3 spalter, og der henvises til 75 termer).

Den saglige tilgang til ordene kan undertiden "forføre" redaktørerne til at skrive rent saglige definitioner, realdefinitioner. Det kan man se et eksempel på i artiklen *græsrode*, hvor *græsrode* er defineret "dss *dugrode*" (bilag 2). Når hørren var høstet, *rødede* man den ved at lægge den i vand. (Verbet *røde* er beslægtet med adjektivet *rådden* og betyder egl. 'få til at rådne'). Hensigten var at mørne taverne, før hørren skulle *brydes* og *skættes*. Man kunne enten *vandrøde* hørren ved at lægge den i et vandhul, eller man kunne *græsrode* eller *dugrode* den ved at lægge ud på græsset, så duggen kunne falde på den. Sagligt set er *græsrode* altså det samme som *dugrode*, men sprogligt set er det noget andet: *græsrode* betyder 'røde ved at lægge ud på *græsset*', mens *dugrode* betyder 'røde ved at lægge ud i

duggen'. Definitionen af *græsrode* er atypisk, idet vi som regel tager hensyn til ordenes (sammensætningernes) sproglige indhold.

Konklusion

Det saglige stof indtager en anden og større plads i ØMO end i ordbøger normalt. Vi føler os forpligtede til at formidle det saglige stof, som findes i vores materiale, og til at oplyse om den materielle kultur, som sproget hører sammen med. Vi overskrider bevidst grænsen mellem ordbog og leksikon, men vi forsøger at erkende og markere grænsen: Vi opererer med særlige leksikonafsnit i vores taksonomi - som i artiklen *drivgarn*. Vi markerer, at citaterne er sagligt beskrivende ved at sætte dem i anførelstegn - som i artiklen *bleg*, i modsætning til replik-citaterne, der er i kursiv. Og i definitionerne sætter vi de rent saglige, leksikonagtige oplysninger i parentes - som i artiklen *blusselygte*.

For at markere ordbogens særpræg har vi valgt at kalde ordbogen "en sproglig-saglig ordbog" (i ordbogens undertitel). Vi har ikke opfundet en ny genre, men vi har arbejdet bevidst med en blandet genre.

Noter

- 1) Jf. Sven Tarp: Encyclopædiske oplysninger i tosprogede tekniske ordbøger, i R.V. Fjeld (udg.): *Nordiske Studier i Leksikografi. Rapport fra Konferanse om leksikografi i Norden. 1992.*

Bilag 1: Artikler fra Ømålsordbogen

drivgarn (også **drivegarn**) s – (fiskerspr) fiskegarn til at *drive* (II.8) med (efter sild mm) (mod *sættegarn*) Ø(alm): "de to tredjeparter af sæsonen det var med *drivegarn* i Bæltet" (resten med *sættegarn*) L(Vindb <Onsevig>), "vi kendte ikke andet end bundgarnsfiskeri med sild til 1888, så var vi et par stykker og fik [*dri·wga·s'n*]" nS(Gill); drivgarnsfiskeri foregik om efteråret (efter *høstsild*) især i Langelandsbæltet, Storebælt og Øresund; bådene var større både (fx *bæltsbåde*); garnene (*sildegarn, næringer, hankegarn*) blev bundet sammen i en *lænke (længde)* der med et *drivtov* blev gjort fast foran i båden; garnlænken og båden drev med strømmen; garnene blev enten *drevet ud* (idet båden drev med vinden ad læ til, mens man →*gav* (1) *garnene ud*) ell *sejlet ud* (idet båden sejlede for små sejl ell motor); garnlænken stod på tværs af strømmen, enten oppe i vandet ell sænket lidt (jf *hankegarn, kobbe v*); man kunne hale det inderste garn (*kølgarnet, stævnnæringen*) ind for at se, om der var sild i det; det dårligste garn blev ofte sat yderst (som *hyr(d)e, rumpegarn, rumpenæring*); jf Drechsel.Saltvfisk:37ff og kort E, DaSaltvandsfiskeri(1935):75.

leksikonafsnit

drivtov (også **drivetov**) s – (fiskerspr) tov hvor med *drivgarnene* gøres fast til båden Am(Drag), nS(Hum <Sletten>), L(Vindb <Onsevig>), Lgl-(Long); jf *drivline, garntov, køltov, liggetov, -trosse*.

Bilag 2: Artikler fra Ømålsordbogen

bleg I s f, dog > c S, c Æ – mht lydformer se *bleg* III –1) det at bleges (om lærred, garn); især i forb som ligge, være, hænge på (ell til) *bleg* Ø(alm): "lærredet *lå på bleg*, der var stropper i siderne med pinde stukket igennem; det blev taget ind om aftenen og lagt i vand og lagt vådt ud om morgenen og flere gange om dagen stænket" nS(Gørls), jf APed.BS:98f, "det lærred der *lå til bleg* blev vandet flere gange i døgnet; der blev syet små sejlgarnsøjer i og pinde stukket i for at holde det" Fa(Sklb); "garnet fik *et par måneders bleg*, idet en stang blev stukket ind foroven og forneden til at holde garnet i stræk" T(Tås); [*de hj:ɜ jO jæ:ɪn* (hed jo gerne) *a vi sge ha ga:ɪnəd pO blæj te fad'læwn*] Lgl(Hu) || hertil smstnr som *bleg(e)dam*, *-eng*, *-have*, *-mand*, ..

sagligt beskrivende citater
i anførselstegn,
fraseologiske del i kursiv

blusselygte s – (fiskerspr) lygte til at *blusse* (2.5) med ved ålefangst (først tran- ell tællelysygte, senere petroleum- ell karbidlygte) sS(VordLd), Fa(NAls), øF(Kert), sF(Eg, Horne <Bøjden), T(Bjb, Land), Lgl(Lu <Ristinge>); jf *blusse-lampe*.

definition indeholdende
saglige oplysninger

græsørde v – dss *dugørde* (hør) vF(Lu).

saglig definition
(real definition)

dugørde v – (til *dug* I) i forb *dugørde* hør *røde* hørren (straks efter ruskningen) ved at brede den tyndt ud på marken og lade den ligge et stykke tid udsat for dug og regn (mods *vandrøde*) nS(Sno), øF(Dlb, Hjb), vF(Dres), sF(Åstr), Lgl(Trk), Æ(Rise); jf *dug* I, *dugblege*, *-brede*, *græsørde*.

Bilag 3: Uddrag af artikel fra Ømålsordbogen

garn s n ...

-2) fiskeredskab ..

|| br i mere eller mindre faste forb m vb (vedr klargøring og vedligeholdelse:) *barke, baske, bøde, flådde, hanke, klove, koge, kviste, lynde, rense, tjære, tørre, stene, øre garn, hænge garn op, klare garn op, skyde garn op, tage garn sammen* mfl og (vedr fangst:) *drage, hale, røgte, trække garn, drive, give, ro garn ud, sætte garn mfl*, jf u de resp vb || i talem: *hun hae sinn ga'n u ætter ham* "arbejdede på at vinde hans gunst" vF-(Kavs), tilsv øS(Str), [*hæn hɔ fad ham i sen-gø'zn*] øS(Str), tilsv L(To Lollo), vF(Ørst) || i folketroen: "Han skulle sætte en Synål i et af Garnets Flaad eller bore Hul og stikke en Pind Flyverøn derind .. det hjalp nogle Gange paa Garnene" nvS(KHels Gravl.ByF:152), "når man ikke havde held med garn, skyldtes det, at de var forheksede; så skulle en kone sent om aftenen gå ned og sætte sig i båden, så ville den indfinde sig som havde forhekset garnene" nS(Gill), "vi måtte ikke møde en gammel kone, når vi gik om med garnene; engang (i 1940erne) kom der en gammel kone, da vi lige havde smidt alle garnene ned, og (hun) vadede hen igennem dem .. og det gik óg rent galt ad helvede til det efterår" Lgl-(Mglb <Bagenkop>), tilsv tro øF(Kert KrF-Sinnet.FiskerlivKert:72), "engang i sildefisket (i 1890erne) havde en båd ikke held med sig, så fandt man en dag en saks i en næring i samme båd" nS(Gill), "hvis ikke de havde nogen sild Mikkelsnat, så var der altså nogen, der havde læst over garnene" Lgl(Mgl <Bagenkop>), "vi måtte ikke fløjte, mens vi stod og lagde garn ned" smst, "det måtte aldrig være et lige tal, det skulle enten være 29 eller 31 garn (i drivlængden)" smst || hertil smstnr som: **bundsætte-**, **driv(e)-**, **fiske-**, **flyde-**, **holde-**, **hytte-**, **kaste-**,

sagligt beskrivende citater,
uden sproglig relevans

Ordboksredaktör i dag

- personliga reflexioner kring en yrkesroll

Inger Hesslin Rider

När mitt första barn föddes i början på sjuttioalet var det ingen tvekan om hur en baby skulle behandlas: mat var fjärde timme, väck barnet om det sover när det är matdags osv. Reglerna var klara och entydiga: så här gör man. Att inte göra så betraktades snarast som hållningslöst och konstigt. Den som blir mamma idag får helt andra instruktioner: det finns inte bara ett sätt, låt situationen avgöra vad som behövs, det är normalt att den ena dagen inte är den andra lik!

Jag har varit mamma längre än jag har varit ordboksredaktör, men under mina femton år i branschen har jag varit med om en väldig utveckling och breddning av redaktörsarbetet. Det finns faktiskt tydliga paralleller till det jag just sagt om spädbarn - inte bara det att det är en lång graviditet och en förlossning som kan vara nog så dramatisk, inte bara det att ordboken känns som ens barn - utan också detta att det förr var mera entydigt hur man skulle göra och att man nu anpassar varje projekt till omständigheterna och ständigt prövar nya vägar.

Min tanke är att beskriva ordboksredaktörens arbete på ett stort affärsdrivande bokförlag. Jag berättar från min egen horisont, på mitt förlag, i Sverige - både som jag upplevde den mera självklara äldre situationen och hur jag ser dagens mycket mera oförutsägbara och skiftande tillvaro. Jag betonar igen att detta är min version, för jag är väl medveten om att det finns skillnader både mellan förlag och mellan länder. Jag hoppas i alla fall att ni ska tycka att det har ett allmänt intresse!

Om man för ett ögonblick går ännu längre tillbaka, så var ju ordböcker förr i allmänhet enmansverk, kända under författarnamnen: Webster, Harlock, Auerbach. Förlagets roll var marginell. Författaren gjorde sitt eget upplägg och arbetade fram en bok, som förlaget gav ut med den textmängd det blev, när författaren ansåg sig vara klar. Någon ordboksredaktion på förlaget behövdes inte. Norstedts ordboksredaktion som jag tillhör, har genom omorganisationer hetat Svenska Bokförlaget, Läromedelsförlagen, Esselte Studium och Esselte Ordbok, men har faktiskt sitt ursprung i *Norstedts korrekturavdelning*. Detta talar sitt tydliga språk om graden och arten av ingrepp i ordbokstexten från förlagets sida.

När jag började hade förlaget sedan länge en fast, författande ordboksredaktion. Jag kom in i en välstrukturerad, väldefinierad och trygg värld - och säkert den bästa skola man kan tänka sig när det gäller det här yrket.

På redaktionen arbetade man i *en* grupp, alla med samma projekt. Alla var väl insatta

i det arbete som pågick och samma regler gällde för alla. Alla resurser satsades på en sak i taget. Det fanns en inbyggd kontinuitet och en stor trygghet i detta. Jag kan inte minnas att det någonsin fanns någon tvekan om hur nästa arbetsdag skulle gestalta sig för ordboksredaktören. Det redaktionella arbetet fyllde i stort sett all ens tid. Jag skulle ljugna om jag nu påstod att jag aldrig har längtat tillbaka till detta. Det som känns som en oerhörd lyx idag: att få sitta en hel lång dag och bara koncentrerat arbeta på sin ordbokstext!

Men man kunde också känna att ordboksmakeri var en syssla som på något vis var isolerad från omvärlden: en skön konst som bara ett fåtal begrep sig på och hade insyn i. Och även om företagsledningen ibland kunde komma med försynta frågor om det inte gick att arbeta *lite* fortare, så behandlades ordboksutgivningen i en kategori för sig på förlaget. Man såg med tålmodigt överseende på den seriösa och flitiga, men något underliga ordboksredaktionen.

Det var en lugn, förutsägbar och kanske lite tråkig värld. Den som arbetar som ordboksredaktör på ett bokförlag idag befinner sig mitt i hetluften: det är en hetsig, rörlig och spännande vardag. Formellt sett är det samma befattning, men allt har egentligen ändrats.

I stället för att ha hela redaktionen knuten till ett och samma projekt arbetar vi nu i projektgrupper. Vi har en mängd frilansare, både svenska och utländska. Mellan projekten måste resurserna fördelas och prioriteringar göras. I stället för gamla gröna skrivmaskiner och bruna kortlådor har vi kraftfulla datorer med lysande blå skärmar. I stället för att på heltid arbeta med textredigering är ordboksredaktören idag projektledare - vilket betyder att han eller hon sysslar med *planering och ekonomi, handledning och regelskrivning, granskning och kontroll, uppmuntran och kritik, logistik och statistik, marknadsföring och produktutveckling och dataarbete i teori och praktik* - samt förstås *lexikografiskt arbete!*

Ordbokstext är fortfarande grunden i arbetet, men ordboksredaktören ägnar sig mer åt att se till att andra gör ett arbete som på olika sätt håller måttet, än åt att själv skriva.

På den gamla ettprojektsredaktionen satt redaktionschefen med de flesta av dessa olika tåtar i sin hand. Utvecklingen av redaktörens roll är förstås inte resultatet av något slags samlad aktion för att göra alla till småchefer utan är en anpassning till situationen och marknaden. Dagens samhälle kräver ordböcker på många språk, aktuella och anpassade till olika användare.

På en redaktion där 20-30 titlar kan vara i arbete samtidigt kan inte redaktionschefen leda och styra allt i detalj.

Gruppen av redaktörer och projektledare måste ha ett nära samarbete. Det gäller att förvalta den samlade kompetens som finns inom gruppen. Det gäller att hjälpas åt att hänga med i vad som händer i den akademiska världen, inom lexikografien, på teknologisidan och så vidare. Eftersom projekten också är så olika sinsemellan så är det viktigt att utbyta erfarenheter, både goda och dåliga, så att kompetensen och effektiviteten hela tiden växer. Mycket kraft ägnas åt samarbete kring ordböckernas utformning. Användarvänlighet är ju bland annat att göra så att användaren känner igen sig mellan olika

böcker och inte behöver lära sig nya system för varje ny ordbok. Eftersom resurserna är ändliga måste projekten samplaneras i tid och ekonomi. En redaktör kan bli tvungen att acceptera nedskärning i sitt projekt under en tid till förmån för ett annat mera akut. Medarbetare kan "lånas ut" mellan projekten vid toppar och dalar. Målet är i alla lägen naturligtvis att få fram bästa möjliga ordbok till rimliga kostnader och inom rätt tid.

För att åstadkomma detta måste dagens ordboksredaktör klara de olika sysslor som jag nyss nämnde.

Planering är A och O i arbetet. Ju fler projekten är, desto viktigare blir detta moment. Resurserna måste användas på bästa sätt. För budgetering måste redaktionschefen få realistiska underlag att utgå ifrån. Att "tid är pengar" är verkligen ett talesätt med stor giltighet när det gäller planläggning av ordboksarbete.

Redan den allra första kalkylen över ett ordboksprojekt måste bygga på så "hårda fakta" som möjligt. Man får börja med att skapa en tydlig bild av den färdiga produkten och det arbete som ska leda fram dit:

Vilken är marknaden?

Hur stor är den?

För vem, dvs hur stor, ska boken vara?

När ska den vara klar?

Vad kan tekniken göra?

Hur många moment behövs i arbetet?

Hur många redaktörstimmar?

Hur många redaktörer?

Så här kan resonemangen gå: Vi vill ha en ordbok för allmänmarknaden, en medelstor ordbok från svenska till språket X. Omfång ca 55 000 ord och fraser eller ca 700 sidor med normal layout. Den bör vara klar på hösten år Y, dvs om drygt tre år. En svensk bas att utgå ifrån har vi.

Erfarenheter från andra projekt hjälper oss att komma fram till att vi bör klara projektet på två vändor: en författande / översättande och en granskande plus en teknisk kontroll, ren korrekturläsning. Jämförelser med andra liknande projekt hjälper oss att beräkna hur många redaktörstimmar projektet kan komma att kräva.

När kalkylen är uppgjord och godkänd och utgivningsbeslutet formellt fattat kan den egentliga projektplaneringen börja. En detaljerad *projektplan* med tidsangivelser och delmoment görs nu upp.

Vi hade alltså ett svenskt underlag att utgå ifrån och behöver inte avsätta speciell tid för ordurvalet. Det är lämpligt att ordboken kommer ut i september år Y - 36 månader från projektstart.

Därför räknar man baklänges. Vi får räkna med tre månader för sättningsförberedelser, sättning, tryck och bind. I juli månad är Sverige stängt, så den månaden får vi räkna bort. Materialet ska alltså vara klart att lämna oss i april samma år. Återstår 32 månader att fördela på det redaktionella. En rimlig fördelning blir kanske att fas 1, författandet / översättandet, får ta 17 månader, fas 2, granskningen, får ta 12 och de sista kon-

trollerna 3 månader. Varje fas bryts sedan ner ytterligare, till spalter per vecka, eller ofta spalter per timme i slutet av projekten.

Projektgruppen formas, genom kontakter, annonser, intervjuer. Medlemmarna i gruppen får binda sig för ett bestämt antal timmar per vecka. Nu är det dags för *handledning och regelskrivning*. Gruppen måste ha tydlig information om förlagets intentioner vad gäller omfång och disposition av ordboken. Det tar alldeles för mycket tid att göra hur som helst först och sedan försöka samordna till ett helt - det har många av oss bittert erfärit. Även om medarbetarna varit med förr och är vana, är regelboken - det som vi hos oss av tradition kallar Koranen - oerhört viktig.

Som alla i den här församlingen vet så finns det massor av sätt att säga samma sak på i en ordbok, och många av dem är lika korrekta, men om de används om varandra blir det inte en bra och helgjuten ordbok.

I början av arbetet behövs redaktören / projektledaren på olika sätt. Kontroller av producerad text behövs för att avslöja olikheter i behandling inom gruppen, och också enskilda medarbetares "egna konstigheter". Med täta projektmöten får man till stånd diskussioner och lösningar av enskilda redigeringsproblem. Allt detta glesas ut successivt när projektet är stadigt på spåret.

I den idealiska projektgruppen arbetar alla som de ska, när de ska och presterar dessutom ett bra material. Som alltid när vi ofullkomliga människor är inblandade är verkligheten ofta en annan. Projektledaren måste kunna fungera som coach, och ge *uppmuntran och kritik*, piskor och morötter. Problemen visar sig ofta röra tider, och därmed också pengar. En medarbetare kanske har en massa andra sysslor också och gör i snitt 5 timmar i veckan i stället för de 20 som planeringen bygger på. En annan gång är det någon som behöver extra semesterpengar och visar sig göra 40 timmar i stället för de beräknade 20. I bästa fall är det arbete som för projektet framåt i samma takt och det finns dessutom pengar avsatta. Men lika ofta kan det vara sådant som spräcker kostnadsbudgeten - det här projektet är ju inte det enda som är på gång. Alla måste följa reglerna för sin del av verksamheten om verksamheten ska gå runt.

Andra problem inom ett projekt kan vara en medarbetare som inte håller måttet vad gäller kvalitet på det utförda arbetet. Det är inte lätt att komma med det beskedet till en medarbetare som kanske dessutom är i stort behov av inkomsten. Alla medarbetarna i ett projekt är individer. De har väldigt olika behov av kontakt, stöd och feedback. En bra projektledare ska kunna dela ut just detta i individuellt anpassade doser. Har man tur har man själv någon där man kan få hämta stöd när det känns tungt - för det gör det förvisso emellanåt.

Logistik och statistik är aspekter av den planering och uppföljning som måste pågå under projektets hela löptid. Har vi den bästa arbetsgången för att föra projektet framåt? Kan några moment slås ihop? Blir det mer effektivt om B går igenom materialet före A? Hur långt har vi kommit i procent av hela materialet? Kommer tidplanen att hålla? Ligger vi just nu på plan? Om inte - varför inte? För enligt Murphys berömda lag är den här typen av avvikelser nästan alltid till det sämre. Vad kan man göra för att komma tillbaka på spåret? Kanske kan någon medarbetare tillfälligt ta på sig lite mer, eller kanske

måste vi helt enkelt arbeta snabbare. I det sammanhanget kan man komma in på funderingar och diskussioner om överkvalitet. Eftersom ordboken så mycket är ens hjärtebarn vill man ju inte se den behäftad med brister. Men realiteten idag är att en ordbok måste bära sina kostnader - ju längre man filar och polerar desto mer kostar det och desto längre dröjer det innan pengar kommer in igen och gör nya projekt möjliga. Det kan vara nog så smärtsamt att inte kunna göra allt man vill och allt man vet att man skulle kunna göra.

Det jag säger kan låta väldigt krasst ekonomiskt - och visst blir det ett ständigt räkande. Men det ligger ju faktiskt ingen motsats i sig mellan god kvalitet och säljbarhet! Och den enda riktigt bra ordboken är trots allt en som är tillgänglig för fler än dem som skrivit den!

Att man hela tiden ska arbeta fram mot en säljbar produkt är naturligtvis en självklarhet på ett förlag. Långt innan ordboken finns i sinnevärlden måste man börja tänka på *marknadsföring*. Det är för sent att börja sälja in boken när den redan är klar i tryck. Ett halvår eller mer innan färdig bok finns ska katalogtexter skrivas, omslagsskisser beställas och underlag för annonsering lämnas. Ordboksredaktören förväntas inte vara någon copywriter, men är ju den som känner ordbokens kvaliteter bäst. Användarna är ofta kunniga och medvetna och kräver handfast information. Förlagets säljare måste få argument så att de förstår att detta är just den ordbok världen väntat på, så att de i sin tur kan övertyga inköpare i bokhandel eller skola.

Produktutveckling är ett annat sätt att trygga ordboksutgivningens framtid. Målgruppsanpassade ordböcker är något vi kommer att få se mer av. Med teknikens hjälp kan man åstadkomma olika varianter ur samma källa, böcker som skiljer sig åt till innehåll, uppbyggnad eller yttre form eller alltihop. Ordböcker har ju egentligen den underbara egenskapen att alla behöver dem. Det vilar på oss alla att göra det bästa av detta faktum.

Datorerna och den nya tekniken finns överallt i ordboksredaktionens värld. Det är svårt att tänka sig en tid när man klarade sig utan dem. Ordboksredaktören behöver inte vara programmerare, men måste kunna göra en kravspecifikation för varje projekt. Alla projekt behöver inte produceras i ett speciellt system för ordboksarbete. Ibland räcker det med en vanlig ordbehandlare, ibland kanske en laserutskrift kan få duga som original - regeln är återigen att det finns fler än ett sätt.

Man måste också vara öppen för nya vägar och nya experiment: vändningar, olika slags motkörningar m.m. En del experiment blir lyckade, andra tämligen plågsamma och katastrofala, men allt samlas ju till en väldig erfarenhetsbank på redaktionen. Redaktörerna har med åren förvärvat en rätt hälsosam oräddhet för datorer. Och tålmodigare slavar än datorerna får man leta efter. För systematiseringsarbete är det fantastiskt att ta datorerna till hjälp för att städa upp bland inkonsekvenser, samordna förkortningslistor, söka efter specialfall - uppgifter som är näst intill omöjliga att göra "för hand". Men visst kan man med datorhjälp också åstadkomma alldeles gräsliga fel - det gäller att ha ögonen med sig. Redaktören har ansvaret för att allt finns på plats där det ska. Det kan inte vara lustigt att vara redaktör för den svensk-nederländska ordbok (inte vår!) där bokstaven S oväntat slutar vid "stadigvarande" och alltså saknar både "stå", "ställa" och "säga" - för att

Inger Hesslin Rider

inte tala om "svensk".

Eget ordboksarbete då? Ja, som jag sa i inledningen är det något önskvärt som nästan känns som en lyx. Det är svårt att få tiden att räcka till för allt som ingår i arbetet. Men det är viktigt att man inte tappar bort sin kompetens när det gäller arbete med ordbokstext. Det gäller att finna tid för att skriva själv och att förkovra sig på olika sätt – t.ex. genom sådana här konferenser. För man blir en bättre instruktör och gör bättre beräkningar om man själv är en bra redaktör, det är jag övertygad om!

Slutligen

För att återanknyta till det där med barnuppfostran och spädbarnsvård: det barn jag nämnde i början är nu 19 år, pigg, frisk och – så vitt jag som hennes mamma kan bedöma – fullt normal. Hon är alltså ett levande bevis på att det gick att göra på det viset.

De ordböcker som gjordes på det "gamla" sättet är också utmärkta ordböcker. Men på samma sätt som det är orimligt att låta ett barn gråta av hunger, eller tvinga det att äta när det bara vill sova, är det orimligt att arbeta långsamt när det går att arbeta fortare, att göra få ordböcker när det går att göra många. Ett brett och aktuellt ordbokssortiment som folk vill köpa och använda ser jag som den bästa garantin för att få fortsätta i det här fantastiskt roliga yrket.

Lexikografins terminologi: systematiska aspekter

Bo Svensén

1. Systematik i fackordböcker

1.1. *Encyklopedisk och terminologisk systematik, begrepps- och termsystematik*

I ett fackordboksprojekt kan man skilja mellan olika typer av systematik. Man kan t.ex. skilja mellan encyklopedisk och terminologisk systematik.

Den encyklopediska systematiken avser fakta om ämnesområdet som sådant, medan den terminologiska avser fackspråket inom ämnesområdet.

Den terminologiska systematiken kan i sin tur indelas i begreppssystematik och term-systematik.

Begreppssystematiken avser innehållssidan. Det gäller här bl.a. att utreda vilka begreppsfält, begreppsrelationer och begreppssystem som förekommer inom fackområdet i fråga och tillämpa de vunna insikterna vid insamling, urval, bearbetning och presentation av data.

Termsystematiken avser uttryckssidan. Den blir aktuell främst i sådana fall där den planerade ordboken eller ordlistan gäller ett ungt fackområde, där terminologin ännu inte har stabiliserat sig och där syftet åtminstone delvis är normativt, eller ett etablerat fackområde vars begreppsbyggnad och därmed terminologi är stadd i omvandling. När det saknas en term för ett visst begrepp – eller när det finns flera termer för samma begrepp – innebär termsystematiken att man om möjligt skapar resp. väljer en term vars form avspeglar begreppets förhållande till angränsande begrepp i systemet.

1.2. *Implicit och explicit systematik*

Om man ser det hela i ett annat perspektiv kan man i ett fackordboksprojekt också skilja mellan implicit systematik och explicit systematik.

Den implicita systematiken är inte direkt synlig för användaren. Den avser metoden vid utarbetandet av ordboken och innebär att ett systematiskt synsätt styr dels urvalet av enheter som skall beskrivas, dels sättet att beskriva dessa enheter.

Den explicita systematiken är däremot direkt synlig för användaren. Den avser egenskaper hos den färdiga produkten och innebär att det i den förekommer vissa informationstyper och en viss struktur.

1.2.1. Implicit systematik

Den implicita systematiken kommer in på flera stadier i arbetet: vid insamling och urval av data och vid bearbetning och presentation av data.

1.2.1.1. Under datainsamlingen, då man inventerar ämnesområdets bestånd av termer och via termerna försöker bilda sig en preliminär uppfattning om vilka begrepp som förekommer där, bör man åtminstone ha klart för sig vilka huvudområden ämnet består av och förvissa sig om att de i stora drag blir intäckta av de källor man inledningsvis väljer att använda.

1.2.1.2. Vid urvalet av data krävs en något mer finmaskig systematik. Man måste göra klart för sig vilka begreppsält som de olika huvudområdena består av för att urvalet av begrepp skall bli representativt för respektive område. Men här blir det naturligtvis fråga om ett växelspel: lika väl som lexikografens kunskap om ämnets systematik styr urvalsarbetet, kan detta i sin tur ge nya insikter som förändrar den bild av systematiken som han hade i utgångsläget. Han kanske härigenom upptäcker luckor i materialet, som tvingar honom att anlita flera källor utöver de ursprungligen valda. Han kan också upptäcka begreppsält som blivit alltför rikligt företrädade och därför behöver reduceras. Selektion och systematisering måste gå hand i hand.

Ett hjälpmedel som med fördel kan tillgripas på det här stadiet är en ämnesklassificering av de utvalda enheterna. De generella klassificeringssystem som redan finns kan då visa sig vara alltför grovmaskiga för att på ett meningsfullt sätt kunna tillämpas på ett begränsat fackområde, och det kan då bli fråga om att skraddarsy ett system för området i fråga.

På det här stadiet kan det också vara dags att avgöra vilka termer som skall representera de valda begreppen, och det är nu som termsystematiken kommer in i bilden.

1.2.1.3. Bearbetning och presentation av data kräver naturligtvis den allra mest finmaskiga systematiken. Det gäller att utreda vilka begrepp, begreppsrelationer och begreppssystem som förekommer och låta den systematik som växer fram genom detta arbete styra utformningen av den information som ges om de enskilda begreppen och om deras relationer till andra begrepp.

Om ordboken är flerspråkig spelar systematiken också en viktig roll i samband med harmoniseringen av de inbegripna språkens termer och begrepp.

1.2.1.4. Den ideala arbetsgången när det gäller den implicita systematiken skulle kunna beskrivas som (a) terminventering, (b) begreppsinventering, (c) begreppsselektion och i samband därmed begreppssystematisering, (d) eventuell begreppsharmonisering, (e) termselektion och i samband därmed termsystematisering och (f) eventuell termharmonisering.

1.2.2. *Explicit systematik*

Den explicita systematiken manifesteras på olika plan i den färdiga produkten: i megastrukturen, i makrostrukturen och i mikrostrukturen.

1.2.2.1. *I megastrukturen*

När ordbokens huvudlista inte är systematiskt ordnad, måste den, om man vill åskådliggöra begreppssystematiken för användarna, kompletteras med andra komponenter. Dessa kan placeras på olika ställen i boken: (a) före huvudlistan, (b) insprängda i huvudlistan utan att för den skull utgöra en del av den och (c) efter huvudlistan. De utgör då, för att använda den terminologi som brukas av Hausmann & Wiegand (1989, s. 330 f) i deras utredning av ordbokens textuella struktur, (a) "front matter", (b) "middle matter" och (c) "back matter"; sammanfattande benämning "outside matter" (på svenska förslagsvis "förtext", "mellantext", "eftertext" och "utanförtext").

1.2.2.1.1. En sådan komponent i utanförtexten är en systematiskt uppbyggd introduktion till ämnesområdet, som i huvudsak förmedlar encyklopedisk information men indirekt också terminologisk. Den kan placeras som förtext eller eftertext. Normalt kan den inte behandla alla begrepp som förekommer i huvudlistan utan får nöja sig med de mest grundläggande. De termer som förekommer i huvudlistan är ofta markerade med avvikande teckensnitt:

§ 4 **Eukaryoter**

Eukaryote organismer indeholder **organeller**. Den største af disse, **cellekernen**, forekommer ofte kun i ét eksemplar pr **celle**. Den er omgivet af en dobbeltlaget, glat **membran**, kernemembranen, og er herved adskilt fra cellens andre bestanddele, **cytoplasmaet** og de øvrige organeller. Cellekernen indeholder alt kerne-DNA, som er organiseret i form af **kromatin**. Der er konstant forbindelse mellem kerne og cytoplasma via et stort antal porer i **kernemembranen**.

De vigtigste blandt de øvrige organeller er **mitokondrierne** (figur 4) og — hos de grønne planter — **kloroplasterne** (figur 5), også kaldet **grønkorn**. Mitokondrier og kloroplaster er ligesom cellekernen omgivet af en dobbeltmembran, dog af en noget anden beskaffenhed end kernemembranen.

Fig. 1 (Kaufmann & Bergenholtz, 1992)

Storleksförhållandet mellan introduktionen och huvudlistan kan växla. I Kaufmann & Bergenholtz (1992) upptar den systematiska översikten 38 av bokens 445 sidor. Som exempel på en mellanform kan nämnas Sponset & al. (1987), där den systematiska delen är placerad som eftertext och omfattar ca 300 av totalt 600 sidor. I andra änden av skalan befinner sig en typ där den systematiska introduktionen till ämnesområdet är huvudsaken och upptar i stort sett hela boken, t.ex. Sjögren (1979).

Behovet av systematiska ämnesöversikter är kanske allra störst när ordboken är två- eller flerspråkig och ämnesområdet dessutom är i hög grad kulturspecifikt, eftersom man då också måste utreda grundläggande skillnader mellan förhållandena inom de olika språkområdena. Ett typiskt sådant ämne är juridiken.

Separation og skilsmisse

Spanien

Separation (*separación -personal- jurídica*) og skilsmisse (*divorcio*) kan kun opnås ved retssag, dvs. ved dom.

Danmark

Separation og skilsmisse kan opnås på to måder, nemlig ved dom eller administrativt, dvs. ved bevilling fra en administrativ myndighed, nemlig (stats-)amtet, også kaldet overøvrigheden. For at der kan gives separation eller skilsmisse ved bevilling, skal parterne imidlertid være enige om at anvende denne fremgangsmåde samt om vilkårene for separation eller skilsmisse.

Fig. 2 (Henriksen, 1991)

1.2.2.1.2. Begreppssystematiken kan i utanförtexen också åskådliggöras genom systematiskt uppbyggda termlistor, bland vilka man kan urskilja två huvudtyper när det gäller graden av systematik. Den ena försöker reproducera det underliggande begreppssystemet.

RESOURCE ALLOCATION

Allocation System
1'4'1.1 GOVERNMENT ALLOCATION
1'4'1.2 MARKET ALLOCATION

Ownership of Resources
1'4'2.1 PUBLIC OWNERSHIP OF RESOURCES
1'4'2.2 PRIVATE OWNERSHIP OF RESOURCES

Economic System
1'4'3.1 PLANNED ECONOMY
1'4'3.2 MARKET ECONOMY
1'4'3.3 MIXED ECONOMY
1'4'3.3'1 Economic Distinction
1'4'3.3'1.1 NONMARKET SECTOR
1'4'3.3'1.2 MARKET SECTOR
1'4'3.3'2 Legal Distinction
1'4'3.3'2.1 PUBLIC SECTOR
1'4'3.3'2.2 PRIVATE SECTOR

Fig. 3 (Toft & Roth, 1990)

Den andra är en blandad typ, som visar en mer eller mindre systematisk uppdelning i delområden, under vilka termerna anförs i alfabetisk ordning .

2 GODS OCH GODSENHETER

2.1 Godstyper

2.1.1 Grundläggande begrepp

bulkods
massgods
partigods
slurry
styckegods

2.1.2 Godstyper efter befordringsätt och typ av godsenheter

expressgods
fraktgods
ilgods
pallgods
returgods
skrymmande gods
spridningsgods
vagnslastgods
volymgods

Fig. 4 (TNC 93, 1992)

1.2.2.1.3. En del fackordböcker har en särskild avdelning med grafiska framställningar av begreppssystemen, till vilka hänvisning görs från huvudlistan.

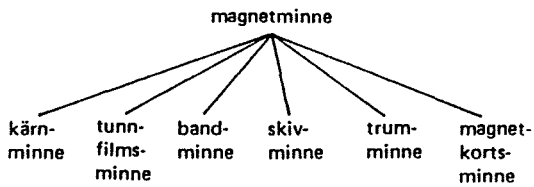


Fig. 5 (SS 01 16 01 / TNC 82, 1984)

1.2.2.1.4. Även termsystematiken kan åskådliggöras i utanförtexten genom olika typer av termlistor. Dessa kan utformas så att flerordstermer och sammansatta termer som har ett visst element gemensamt insorteras på två ställen, dels efter den fullständiga termen, dels efter det gemensamma termelementet

gångstål 3.2.30
 cirkulärt ~ 3.2.33, 29.2.52
 flertandat ~ 3.2.34
 inner- 3.2.31, 29.2.50
 invändigt ~ 3.2.31
 skiv- 3.2.33
 utvändigt ~ 3.2.32
 ytter- 3.2.32, 29.2.51

Fig. 6 (SMS 508 / TNC 78, 1985)

Registret kan också utformas som en KWIC-konkordans.

| | | |
|---------------------------------|----------------|----|
| | oriented graph | 32 |
| complete | oriented graph | 32 |
| connected | oriented graph | 42 |
| converse | oriented graph | 47 |
| degree of a regular | oriented graph | 34 |
| minimally strongly connected | oriented graph | 36 |
| regular | oriented graph | 34 |
| strictly complete | oriented graph | 32 |
| strictly unilaterally connected | oriented graph | 36 |
| strongly connected | oriented graph | 36 |
| transitive | oriented graph | 38 |
| unilaterally connected | oriented graph | 36 |

Fig. 7 (Gavare, 1972)

1.2.2.2. I makrostrukturen

Begreppssystematik kan i makrostrukturen bara visas på ett sätt, nämligen i artiklarnas inbördes ordning i huvudlistan.

1.2.2.2.1. Ordboken kan sträva efter att i makrostrukturen reproducera de underliggande begreppssystemen så långt som möjligt. Systematiken åskådliggörs i regel genom sät-tet att numrera termerna

3.5.3.4

NO fluidvolum
DE Druckmittelverbrauch (m)
EN fluid volume
FR volume (m) de fluide utilisé
SV fluidvolym

3.5.3.4.1

NO væskevolum
DE Hydrovolumen (n)
EN hydraulic fluid volume
FR volume (m) de fluide hydraulique
SV vätskevolym
Bestemmes ved rådende atmosfærisk trykk og omgivelsestemperatur.

3.5.3.4.1.1

NO teoretisk væskevolum, geometrisk slagvolum
DE theoretisches Hydrovolumen (n)
EN theoretical oil volumes, geometric cylinder capacity
FR volume (m) théorique de fluide hydraulique, cylindrée (f)
géométrique du vérin
SV ideal vätskevolym
Beregnet væskevolum for én arbeidsyklus.

Fig. 8 (RTT 56, 1988)

1.2.2.2.2. Ofta är det inte meningsfullt att försöka återge några begreppssystem; det kan ju faktiskt vara så att det inte finns några klart urskiljbara sådana att återge. Då brukar man åtminstone försöka upprätta en hierarki av begreppsfält, som får fungera som rubriker under vilka man sedan presenterar dithörande termer i alfabetisk ordning.

4.3 Coagulation phenomena

| No. | Term | Definition |
|------|---------------|---|
| 4301 | agglomeration | Reversible or irreversible joining together of latex particles. |
| 4302 | blowing down | Removal of excess ammonia from latex by stirring the latex while passing a stream of air across the surface. |
| 4303 | coagulation | Irreversible agglomeration of particles originally dispersed in a rubber latex to form a continuous phase of the rubber and a dispersed phase of the serum. |
| 4304 | coagulum | Immediate product of coagulation of latex consisting of agglomerates of rubber particles. |

Fig. 9 (BS 3558, 1980)

1.2.2.3. I mikrostrukturen

Den information som ges i mikrostrukturen och som kan bidra till att åskådliggöra systematiken är av flera slag.

1.2.2.3.1. Den encyklopediska systematiken kan givetvis tillgodoses genom en särskild avdelning inom artikeln som ger sakinformation.

1.2.2.3.2. Begreppssystematiken kan åskådliggöras exempelvis genom att ett antal begrepp behandlas under sitt gemensamma överbegrepp .

sluss

anläggning avsedd att möjliggöra genomfart av fartyg mellan olika vattennivåer i en farled

Slussen kan bestå antingen endast av en slussport (*enkelsluss*) eller av en bassäng med slussport i båda ändarna (*dubbelsluss*). En grupp av flera slussportar i en kanal kallas *slusstrappa*. Sluss eller slusstrappa ersätts ibland av en anordning bestående av ett vertikalt rörligt eller på lutande spår åkbart tråg, benämnt *fartygshiss*.

en lock

fr *écluse* f

de *Schleuse* f

Fig. 10 (TNC 93, 1992)

I tvåspråkiga ordlistor som bara ger ekvivalenter på målspråket kan det ibland också vara till hjälp för systematiken om sammansatta termer och flerordstermer där ett visst ord utgör mellan- eller slutled inte bara anförs på sin alfabetiska plats utan också sammanförs under det gemensamma ordet.

avfall waste [61. 62], discharge [E]; **animaliskt** ~ animal waste [62]; **blandat** ~ mixed waste [62]; **brännbart** ~ combustible waste [62]; **bygg** ~ construction waste [62]; **fast** ~ solid waste [62]; **flytande** ~ liquid waste [62]; **förpacknings** ~ packaging waste [62]; **gasformigt** ~ gaseous waste [62]; **gatu** ~ road sweepings (pl.) [62]; **hushålls** ~ domestic waste [61. 62]; **hygges** ~ tree trimmings (pl.)

Fig 11 (TNC 91, 1990)

Det förekommer också att man nöjer sig med att under en viss term räkna upp ett antal underordnade termer med genomskinlig betydelse utan att nödvändigtvis förse dem med definitioner, vare sig här eller på deras alfabetiska plats.

färja

fartyg speciellt utformat för transport av passagerare, bilar, tåg o d, vanligen i anslutning till landledd

Med förled specificeras lasttyp:
bilfärja, godsfärja, passagerarfärja,
tågfärja.

en ferry

fr *bac* m

de *Fähre* f, *Fährschiff* n

Fig. 12 (TNC 93, 1992)

1.2.2.3.3. Systematiken kan också åskådliggöras genom explicita hänvisningar, exempelvis till andra enheter i makrostrukturen; en sådan hänvisning avser som regel sidordnade (se Fig. 13 a) eller underordnade begrepp (se Fig. 13 b).

vindgradient

gradient som anger ändring i vindhastighet i den riktning där ändringen är störst; jfr *temperaturgradient*

- E wind gradient
- F gradient m de vent
- D Windgradient m

gradient

en storhets förändring genom längd i den riktning där förändringen är störst; jfr *temperaturgradient, vindgradient*

- E gradient
- F gradient m
- D Gradient m

Fig. 13 a och b (TNC 81, 1984)

Hänvisningar kan också gå till enheter i utanförtextern som ger systematisk information, t.ex. till ett visst kapitel eller en viss paragraf i introduktionen till ämnet (se Fig. 14) eller till en viss figur i en särskild avdelning med grafiska framställningar av de olika begreppssystemen (se Fig. 15).

genome-walking *genome walking/chromosome walking*
= genomvandring
→ kromosomvandring; § 24

magnetkortsminne →7

minne som utnyttjar *magnetkort*

- E magnetic card storage
- F mémoire f à cartes magnétiques
- D Magnetkartenspeicher m

Fig. 14 (Kaufmann & Bergenholtz, 1992)

Fig. 15 (SS 01 16 01 / TNC 82, 1984)

1.2.2.3.4. En klassningskod, som visar begreppets plats i systematiken eller åtminstone anger begreppsältet, kan också vara till en viss hjälp.

1.2.2.3.5. Termnummer är ytterligare ett sätt att åskådliggöra begreppets plats i systemet; sådana nummer kan också användas vid implicita hänvisningar inuti själva definitionerna.

6.2 Présentation

6.2.1 **dictionnaire**: Répertoire structuré d'unités lexicales comportant des informations linguistiques sur chacune d'entre elles.

6.2.1.1 **dictionnaire terminologique** (terme toléré: **dictionnaire technique**): Dictionnaire (6.2.1) qui comprend des données terminologiques (6.1.5) relatives à un ou plusieurs domaines (2.2) particuliers.

6.2.1.1.1 **vocabulaire**: Dictionnaire terminologique (6.2.1.1), basé sur un travail terminologique (8.2), qui présente la terminologie (5.1) d'un domaine (2.2) particulier ou de domaines (2.2) associés.

Fig. 16 (ISO 1087, 1990)

2. Systematik inom fackområdet lexikografi

2.1. Encyklopedisk systematik

2.1.1. Systematik i föreliggande framställningar av ämnet

Om man vill bilda sig en uppfattning om den encyklopediska systematiken inom ämnet lexikografi, faller det sig naturligt att undersöka dispositionen i några av de mer eller mindre omfattningsrika och mer eller mindre heltäckande framställningar av ämnet lexikografi som föreligger.

Det verk som det då ligger närmast till hands att undersöka är givetvis *Wörterbücher* (1989–91) [Fig. 17]. Siffrorna i uppställningen avspeglar utgivarnas disposition; de versala bokstäverna visar hur man skulle kunna sammanföra de olika avdelningarna till större block. Grovt kan man där urskilja sju sådana block. Block A behandlar lexikografen i ett närmast sociologiskt perspektiv, B går in på allmänna aspekter av ämnets historia och teori och D tar upp allmänna metodfrågor. Block C ägnas helt åt enspråkiga ordböckers teori och typologi, medan block F uteslutande tar sikte på motsvarande aspekter av den tvåspråkiga lexikografen. Blocken E och G är kanske inte fullt lika intressanta i det aktuella sammanhanget.

A

1. Lexikographie und Gesellschaft
 - 1.1. Wörterbücher und Öffentlichkeit
 - 1.2. Wörterbücher und ihre Benutzer

B

2. Geschichte und Theorie der Lexikographie: Allgemeine Aspekte

C

3. Theorie der einsprachigen Lexikographie
 - 3.1. Bauteile und Strukturen von Wörterbüchern
 - 3.2. Ausgewählte Beschreibungsprobleme im allgemeinen einsprachigen Wörterbuch
4. Wörterbuchtypen
 - 4.1. Allgemeine Aspekte der Wörterbuchtypologie und allgemeine einsprachige Wörterbücher
 - 4.2. Syntagmatische Spezialwörterbücher
 - 4.3. Paradigmatische Spezialwörterbücher
 - ...
 - 4.10. Auf Texte bezogene Wörterbücher

D

5. Arbeitsverfahren in der Lexikographie

E

6. Lexikographie der Einzelsprachen

F

7. Theorie der zwei- und mehrsprachigen Lexikographie
 - 7.1. Prinzipien und Bauteile
 - 7.2. Ausgewählte Beschreibungsprobleme
8. Typologie und ausgewählte Typen der zwei- und mehrsprachigen Lexikographie
9. Die zweisprachigen Wörterbücher in Geschichte und Gegenwart

G

10. Lexikographie von Hilfssprachen und anderen Kommunikationssystemen

Fig. 17. Dispositionen i Wörterbücher 1989–91.

De övriga allmänna framställningar som jag har jämfört med (Zgusta, 1971; Dubois & Dubois, 1971; Kipfer, 1984; Landau, 1984; Hausmann, 1985) är inte till särskilt stor hjälp, främst därför att viktiga komponenter saknas i dem. Kipfer och Landau ignorerar på anglosaxiskt vis de tvåspråkiga ordböckerna helt och hållet; Kipfers framställning är i övrigt systematisk men något naivistisk, Landaus mycket läsvärd men långt ifrån systematisk (jfr Svensén, 1992). Dubois & Dubois tilldelar de tvåspråkiga ordböckerna 5 sidor. Zgusta ägnar 200 av sina 350 sidor snarare åt lexikologi än åt lexikografi. Hausmann är systematisk och koncentrerad men alldeles för kortfattad.

Mitt eget bidrag till genren (Svensén, 1987) skiljer sig något från de andra därigenom att indelningen i enspråkiga och tvåspråkiga ordböcker görs på kapitelnivå och enbart avser själva betydelsebeskrivningen. När det gäller övriga informationstyper görs indelningen på en ännu lägre nivå, nämligen i underavdelningar av kapitel.

2.1.2. En tänkbar systematik

En encyklopedisk systematik över ämnet kan med fördel huvudsakligen bygga på *Wörterbücher* (1989–91), modifierad bl.a. genom att man upptar vissa drag från de andra.

1. Lexikografi och samhälle
 - 1.1. Ordboken och allmänheten
 - 1.2. Ordboken och dess syften/användning
 - 1.3. Ordbokskritik
2. Lexikografin och angränsande discipliner
3. Lexikografins historia
4. Allmän lexikografisk teori
5. Enspråkiga allmänordböcker
 - 5.1. Allmänt
 - 5.2. Typologi
 - 5.3. Strukturer
 - 5.3.1. Megastruktur
 - 5.3.2. Makrostruktur
 - 5.3.3. Hänvisningsstruktur
 - 5.3.4. Polysemistruktur
 - 5.3.5. Mikrostruktur
 - 5.4. Informationstyper
 - 5.4.1. Identifikation
 - 5.4.2. Betydelsebeskrivning
 - 5.4.3. Syntagmatisk information
 - 5.4.4. Paradigmatisk information
 - 5.4.5. Markeringar
 - 5.4.6. Kommentarer

- 5.4.7. Illustrationer
- 5.4.8. ...
- 6. Tvåspråkiga allmänordböcker
 - 6.1. Allmänt
 - 6.2. Typologi
 - 6.3. Strukturer
 - 6.3.1. Megastruktur
 - 6.3.2. Makrostruktur
 - 6.3.3. Hänvisningsstruktur
 - 6.3.3. Polysemistruktur
 - 6.3.4. Mikrostruktur
 - 6.4. Informationstyper
 - 6.4.1. Identifikation
 - 6.4.2. Ekvivalenter
 - 6.4.3. Betydelsedifferentiering
 - 6.4.4. Syntagmatisk information
 - 6.4.5. Markeringar
 - 6.4.6. ...
- 7. Specialordböcker
 - 7.1. Allmänt
 - 7.2. Typologi
 - 7.3. Syntagmatiska specialordböcker
 - 7.4. Paradigmatiska specialordböcker
 - 7.5. Ordböcker över vissa lemmatyper
 - 7.6. Ordböcker över vissa informationstyper
 - 7.7. Ordböcker över vissa språkvarianter
 - 7.8. Ordböcker över vissa texter
 - 7.9. Ordböcker för inlärning
 - 7.10. ...
- 8. Lexikografiska metoder
 - 8.1. Insamling och urval av data
 - 8.2. Bearbetning av data
 - 8.3. Presentation och distribution av data
 - 8.4. Lexikografiskt projektarbete
 - 8.5. Lexikografiska hjälpmedel
- 9. De enskilda språkens lexikografi

Fig. 18. En tänkbar encyklopedisk systematik

2.2. Terminologisk systematik

På vad sätt skulle nu en terminologisk systematik över ämnet lexikografi skilja sig från den encyklopediska? I många stycken kan nog de två anses överensstämma, men det finns enligt min uppfattning också vissa skillnader.

2.2.1. Det finns element som självklart ingår i den ena typen av systematik men som inte

är lika självskrivna i den andra. Dit skulle jag vilja räkna dels lexikografins historia, dels det som i *Wörterbücher* kallas "Lexikographie der Einzelsprachen", som naturligtvis ingår i en fullständig beskrivning av ämnet men som sannolikt inte bidrar till ämnets terminologiska systematik. De saknas alltså i mitt förslag [Fig. 19]. Omvänt har jag i den terminologiska systematiken infört en särskild avdelning som inte finns i den encyklopediska, nämligen nummer 3 "Lexikaliska enheter och deras egenskaper", som närmast är avsedd att hysa de allmänt lexikologiska termer som man av olika skäl kan anse sig böra ha med i en ordbok över lexikografins termer.

1. Lexikografi och samhälle
 - 1.1. Ordboken och allmänheten
 - 1.2. Ordboken och dess syften/användning
 - 1.3. Ordbokskritik
2. Lexikografien och angränsande discipliner
3. Lexikaliska enheter och deras egenskaper
4. Allmän lexikografisk teori
 - 4.1. Typologi
 - 4.2. Strukturer
 - 4.2.1. Megastruktur
 - 4.2.2. Makrostruktur
 - 4.2.3. Hänvisningsstruktur
 - 4.2.4. Polysemistruktur
 - 4.2.5. Mikrostruktur
 - 4.3. Informationstyper
 - 4.3.1. Identifikation
 - 4.3.2. Betydelsebeskrivning: allmänna aspekter
 - 4.3.3. Syntagmatisk information
 - 4.3.4. Paradigmatisk information
 - 4.3.5. Markeringar
 - 4.3.6. Kommentarer
 - 4.3.7. Illustrationer
 - 4.3.8. ...
5. De enspråkiga allmänordbäckernas särskilda teori och problematik
 - 5.1. Betydelsebeskrivning: definitioner, parafrafer
 - 5.2. ...
6. De tvåspråkiga allmänordbäckernas särskilda teori och problematik
 - 6.1. Betydelsebeskrivning: ekvivalenter
 - 6.2. Betydelsedifferentiering
 - 6.3. ...
7. Specialordbäckernas särskilda teori och problematik
 - 7.1. Syntagmatiska specialordböcker
 - 7.2. Paradigmatiska specialordböcker
 - 7.3. Ordböcker över vissa lemmatyper
 - 7.4. Ordböcker över vissa informationstyper
 - 7.5. Ordböcker över vissa språkvarianter

- 7.6. Ordböcker över vissa texter
- 7.7. Ordböcker för inläring
- 7.8. ...
- 8. Lexikografiska metoder
 - 8.1. Insamling och urval av data
 - 8.2. Bearbetning av data
 - 8.3. Presentation och distribution av data
 - 8.4. Lexikografiskt projektarbete
 - 8.5. Lexikografiska hjälpmedel

Fig. 19. En tänkbar terminologisk systematik

2.2.2. För det andra finns det element som är gemensamma för de båda typerna av systematik men som av olika skäl bör grupperas på olika sätt. Ett sådant fall är uppdelningen i enspråkiga och tvåspråkiga ordböcker, som enligt min uppfattning är helt motiverad när man beskriver ämnet men som inte i alla avseenden är motiverad från terminologisk synpunkt. Många problem och termer är gemensamma, och en uppdelning är väl av terminologiskt intresse främst när det gäller betydelsebeskrivningen och vad därtill hör. I förslaget till terminologisk systematik har jag därför på försök infört en avdelning med allmän lexikografisk teori och sedan inrättat två avdelningar för sådant som är speciellt för enspråkig resp. tvåspråkig lexikografi. Det är t.o.m. möjligt att inte ens dessa specialavdelningar är nödvändiga utan att avdelningen "Betydelsebeskrivning" (4.3.2) skulle kunna inrymma båda.

3. Systematik i *Nordisk lexikografisk ordbok*

I projektplanen förutsätts det, att *Nordisk lexikografisk ordbok* skall vara en "allordbok" eller "allbok" i Wiegands mening (Wiegand, 1988) i så måtto att den skall förmedla dels encyklopedisk information om ämnesområdet, dels information om det fackspråk som används inom ämnesområdet. En sådan ordbok bör normalt utarbetas i samarbete mellan ämnesexperter och lexikografer/terminologer; projektet *Nordisk lexikografisk ordbok* är på den här punkten något speciellt genom att ämnesspecialisten och lexikografen är samma person.

3.1. Encyklopedisk systematik

En av ordbokens tre primärfunktioner är att förmedla fackkunskaper inom ämnesområdet, och det är framför allt här som den encyklopediska systematiken kommer in i bilden. Meningen är att den skall bli tillgodosedd genom att den alfabetiska huvudlistan kompletteras med en förtext bestående av en systematisk introduktion till lexikografien. Indirekt stöder introduktionen också en annan av ordbokens tre primärfunktioner, nämligen produktion på modersmålet.

Den systematiska introduktionen kan varken bli omfångsrik eller detaljerad, och det

blir därför nödvändigt att koncentrera sig på det väsentligaste. Av de olika avdelningarna inom den encyklopediska systematiken [Fig. 18] bör man därför enligt min mening lämna 3 och 9 nästan helt åt sidan, beröra 1 och 2 ganska flyktigt, inte fördjupa sig alltför mycket i 7 och 8 och koncentrera sig på 4, 5 och 6.

3.2. Terminologisk systematik

3.2.1. Explicit systematik

Ordbokens huvudlista är tänkt att vara alfabetiskt ordnad. Detta arrangemang tillgodoser en av ordbokens tre primärfunktioner, nämligen reception på modersmålet. Dessutom tillgodoser det i stort sett alla de fyra sekundärfunktionerna, utom möjligen produktion på främmande språk.

I och med att makrostrukturen skall vara alfabetisk ställs det också krav på den explicita terminologiska systematiken. Denna bör främst komma till uttryck i mikrostrukturen och då i första hand genom hänvisningar, dels till besläktade begrepp, dels till synonymer, dels till den systematiska introduktionen. En annan möjlighet som har diskuterats inom projektet är att också när så är lämpligt behandla ett antal sinsemellan sidordnade begrepp under det gemensamma överbegreppet.

Med tanke på ordbokens två produktionsfunktioner (varav den ena hör till primärfunktionerna), och med tanke på att den systematiska introduktionen måste bli ganska kortfattad, bör man förmodligen också kosta på sig systematiska termlistor i utanförtexten. Däremot torde det knappast vara meningsfullt med grafiska framställningar av begreppssystem etc.

3.2.2. Implicit systematik

När det gäller den implicita systematiken skall jag inskränka mig till att i huvudsak diskutera momenten insamling och urval av data.

Insamlingen av data tog sin utgångspunkt i det trespråkiga registret till *Wörterbücher* (1989–91). Detta register är från terminologisk synpunkt inte helt idealiskt som källa.

För det första utgör det ju ingen autentisk textkorpus utan består av ett urval ur en sådan.

För det andra har detta urval gjorts av respektive bidragsförfattare, och skillnaderna i fråga om terminologisk medvetenhet verkar vara ganska stora.

För det tredje är ju registret egentligen inte främst avsett som ett termregister utan som ett sakregister (även om uppslagsorden i det nästan genomgående utgörs av termer). Detta har bl.a. fått till följd att det saknas en hel del termer som förekommer i de olika bidragen och som i högsta grad tillhör det lexikografiska fackspråket, framför allt s.k. "verkstadstermer", som författarna använder tämligen ofreflekterat. För att råda bot på den bristen har registrets termförråd kompletterats med material från andra källor, men frågan är om det räcker.

För det fjärde finns det en risk för att registrets representativitet är något ojämn när det gäller lexikografins fackspråk i allmänhet. De tunga teoretiska avsnitten har till över-

vägande delen skrivits av två personer, Wiegand och Hausmann; en recensent (Görlach, 1991) har räknat ut att Wiegand ensam har skrivit nästan 16 procent av all text i de två första delarna och att Hausmann har bidragit med 35 av hela verkets totalt 350 artiklar. Dessa två – och framför allt kanske Wiegand – har skapat en egen lexikografisk teori och i samband därmed också delvis en egen terminologi. Och många av de termer de introducerar i exempelvis ett tyskspråkigt bidrag har de också försett med engelska och franska ekvivalenter, vilka samtliga, såvitt jag har kunnat finna, har införts i registret. Detta innebär att vi i registret har en viss övervikt för en begreppsbyggnad och en terminologi som är delvis nya och ännu inte har vunnit burskap överallt.

Min slutsats av detta är att projektet, när detta skrivs, på sin höjd kan anses ha avverkat momenten "termininventering" och "begreppsinventering" (se 1.2.1.4). Ännu återstår det verkliga selektionsarbetet, och som ett viktigt led i detta ingår en granskning och bearbetning av de insamlade termerna och begreppen från systematisk synpunkt. En förutsättning för att detta systematiseringsarbete skall bli framgångsrikt är enligt min mening att man också genomför åtminstone en grov klassificering av samtliga enheter i materialet.

Det kommer också att genomföras en kategoriklyvning bland termer och begrepp. De mer eller mindre "centralt lexikografiska" enheterna skall behandlas mer utförligt; de skall ha en normativ prägel och tillgodose både den terminologiska funktionen och översättningsfunktionen. De mer "perifer lexikografiska" skall behandlas mer knapphändigt; de skall ha en deskriptiv prägel och enbart tillgodose översättningsfunktionen. Som en direkt följd av detta kommer den förra gruppen att förses med både definitioner och ekvivalenter, den senare med enbart ekvivalenter.

Som jag ser det, kommer den terminologiska systematiken att bli ett ofrånkomligt hjälpmedel när det gäller att genomföra kategoriklyvningen mellan centrala och perifera termer och begrepp. Dessutom måste systematiken ligga till grund för bearbetningen av de enheter som klassats som centrala, medan kraven på systematik kanske inte behöver vara lika stora när det gäller de perifera.

Litteratur

- BS 3558 (1980): *Glossary of rubber terms*. London: British Standards Institution.
- Dubois, J. & Dubois, C. (1971): *Introduction à la lexicographie. Le dictionnaire*. Paris: Larousse.
- Gavare, R. (1972): *Graph Description of Linguistic Structures. A Linguistic Approach to the Theory of Graphs*. Stockholm: Almqvist & Wiksell. *Data linguistica*, 5.
- Görlach, M. (1991): [Recension av] *Wörterbücher, Dictionaries, Dictionnaires*. Ein internationales Handbuch zur Lexikographie ..., 1-2, red. F.J. Hausmann & al. (*Dictionaries. Journal of The Dictionary Society of North America*, 13, s. 127-135).
- Hausmann, F.J. (1985): *Lexikographie*. I: C. Schwarze & D. Wunderlich (red.): *Handbuch der Lexikologie* (Königstein: Athenäum), s. 367-411.
- & Wiegand, H.E. (1989): Component parts and structures of general monolingual

- dictionaries: A survey. I: *Wörterbücher ...* 1, s. 328–360.
- Henriksen, T. (1991): *Juridisk ordbog spansk-dansk*. København: Gad.
- ISO 1087 (1990): *Terminology – Vocabulary*. Genève: International Standards Organization.
- Kaufmann, U. & Bergenholtz, H. (1992): *Genteknologisk ordbog. Dansk-engelsk/engelsk-dansk molekylærbiologi og DNA-teknologi*. København: Gad.
- Kipfer, B.A. (1984): *Workbook on Lexicography. A course for dictionary users with a glossary of English lexicographical terms*. Exeter: The University. Exeter Linguistic Studies, 8.
- Landau, S.I. (1984): *Dictionaries. The art and craft of lexicography*. New York: Scribner.
- RTT 56 (1988): *Ordbok for fluidteknikk*. Trondheim: Tapir & Rådet for teknisk terminologi.
- Sjögren, P.A. (1979): *Termer i allmän språkvetenskap. Ett systematiskt lexikon*. Stockholm: Akademilitteratur.
- SMS 508/TNC 78 (1985): *Verkstadsteknisk ordlista. Skärande bearbetning*. Stockholm: Sveriges Mekanstandardisering & Tekniska nomenklaturcentralen.
- Sponsel, K., Wallenfang, W.O. & Waldau, I. (1987): *Lexikon der Anstrich-Technik. 1: Grundlagen*. 8. Aufl. München: Callwey.
- SS 01 16 01/TNC 82 (1984): *Dataordboken*. Utgåva 3. Stockholm: SIS – Standardiseringskommissionen i Sverige.
- Svensén, B. (1987): *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. Stockholm: Esselte Studium & Tekniska nomenklaturcentralen.
- (1992): [Recension av] S. I. Landau: *Dictionaries. The art and craft of lexicography (International Journal of Lexicography, 5:1, s. 79–83)*.
- TNC 81 (1984): *Energiordlista*. Stockholm: Tekniska nomenklaturcentralen.
- TNC 91 (1990): *Svensk-engelsk byggordbok*. Solna: Tekniska nomenklaturcentralen.
- TNC 93 (1992): *Transportordlista*. Solna: SIS-STG & Tekniska nomenklaturcentralen.
- Toft, B. & Roth, E. (1990): *Mikroøkonomisk glosar – samt teoretiske og praktiske betragtninger over terminologi [sic] og videnrepræsentation*. Gjellerup: Systime.
- Wiegand, H.E. (1988): Was eigentlich ist Fachlexikographie? Mit Hinweisen zum Verhältnis von sprachlichem und enzyklopädischem Wissen. I: H.H. Munske, P. von Polenz, O. Reichmann & R. Hildebrandt (red.): *Deutscher Wortschatz. Lexikologische Studien. Ludwig Erich Schmitt zum 80. Geburtstag von seinen Marburger Schülern* (Berlin & New York: de Gruyter), s. 729–790.
- Wörterbücher, Dictionaries, Dictionnaires. Ein internationales Handbuch zur Lexikographie ...* (1: 1989, 2: 1990, 3: 1991), red. F.J. Hausmann, O. Reichmann, H.E. Wiegand & L. Zgusta. Berlin & New York: de Gruyter.
- Zgusta, L. (1971): *Manual of lexicography*. The Hague: Mouton.

Polysemi hos svenska verb

Exemplet *slå* och andra fysiska kontaktverb

Åke Viberg

1. Frågeställningen i ett nötskal

I olika tidigare uppsatser (Viberg 1981, 1992a) har jag analyserat delar av svenskans verbordförråd med utgångspunkt från en indelning i semantiska fält. De mest grundläggande verben inom ett fält är normalt höggradigt polysema. Att systematiskt kunna redogöra för denna flertydighet utgör det utan jämförelse svåraste men samtidigt mest fascinerande problemet vid beskrivandet av verben. Som en illustration kommer jag att använda mig av verbet *slå* och andra verb som hör till fältet fysisk kontakt. Analysen utgör en vidareutveckling av min tidigare analys av detta fält som främst berör differentieringen mellan verben med utgångspunkt från deras grundbetydelse (Viberg 1984).

Ett fysiskt kontaktverb anger typiskt en kroppsrörelse som leder till att en kroppsdel kommer i kontakt med något: *Per slog/sparkade till bollen*. Ofta underförstås kroppsdelen som i detta exempel (*slå* - hand eller något som hålls i handen, *sparka* - fot). De verb som hör till fältet förekommer i ett stort antal karakteristiska betydelsemönster, som samspelar med ett antal likaså karakteristiska syntaktiska mönster (eller syntaktiska ramar). Ett sätt att beskriva denna variation är att utgå från ett grundläggande eller prototypiskt mönster och att sedan försöka relatera de övriga mönstren till detta. Valet av det prototypiska mönstret är delvis intuitivt grundat men motiveras fortlöpande genom att man försöker att visa hur övriga mönster på ett naturligt sätt kan härledas ur det prototypiska. Följande exempel illustrerar den prototypiska användningen av ett antal fysiska kontaktverb:

Per *slog* Pål i magen.

Per *sparkade* Pål i baken.

Per *trampade* Pål på tårna.

Per *träffade* Pål i baken med en välriktad spark.

Per *dunkade* Pål i ryggen.

Per *kittlade* Pål på halsen med ett grässtrå.

Per *strök* lille Pål över håret.

Alla exempel kännetecknas av en gemensam **syntaktisk ram**, som består av en person som subjekt och som objekt. Kasusgrammatiskt fungerar subjektet som Agent och objektet som Erfarare. Frivilligt kan två prepositionsfraser förekomma, varav den första som i exemplen ovan anger vilken kroppsdel hos objektet som träffas. Den andra pre-

positionsfrasen specificerar instrumentet:

NP^Z ___ NP^W (prep NP^Y) (med NP^X)

Exempel: Per^Z slog Pål^W (i magen^Y) (med knytnäven^X/en käpp^X)

Detta syntaktiska mönster är så starkt förknippat med de fysiska kontaktverben att det tillfälligt kan locka fram tolkningen fysisk kontakt även hos verb som hör till helt andra fält: *Per gav Pål på nöten.*

Semantiskt delar de fysiska kontaktverben en komponent som informellt kan beskrivas som att det uppstår en kontakt mellan ett konkret föremål som fungerar som instrument (x) och ett annat konkret föremål (y) som i exemplen ovan utgörs av en kroppsdel hos objektet. Formaliserat kan detta återges: HAPPEN(CONTACT(x,y)) 'det uppstår kontakt mellan x och y'.

En förutsättning för att de givna exemplen uppfattas som prototypiska är att verbet har en bokstavlig tolkning. De beskriver en fysisk kontakt som är direkt observerbar med våra sinnen. Som observatörer av händelserna kan vi se och i vissa fall även höra vad som sker. Genom tidigare erfarenhet kan vi också sätta oss in i objektets situation och vet hur resultatet av de olika handlingarna uppfattas med känslan. Perceptuellt grundade betydelser är i allmänhet primära i förstaspråksinläringen och lättare för barn att lära in. I följande exempel har verben antagit mer abstrakta betydelser ('bese-gra', 'avskeda', 'vålla obehag', 'möta', 'avlägsna'):

Per *slog* Pål i schack.

Per *sparkade* Pål från posten som generaldirektör.

Per *trampade* Pål på tårna genom att påminna om hans fadäs.

Per *träffade* Pål i Köpenhamn.

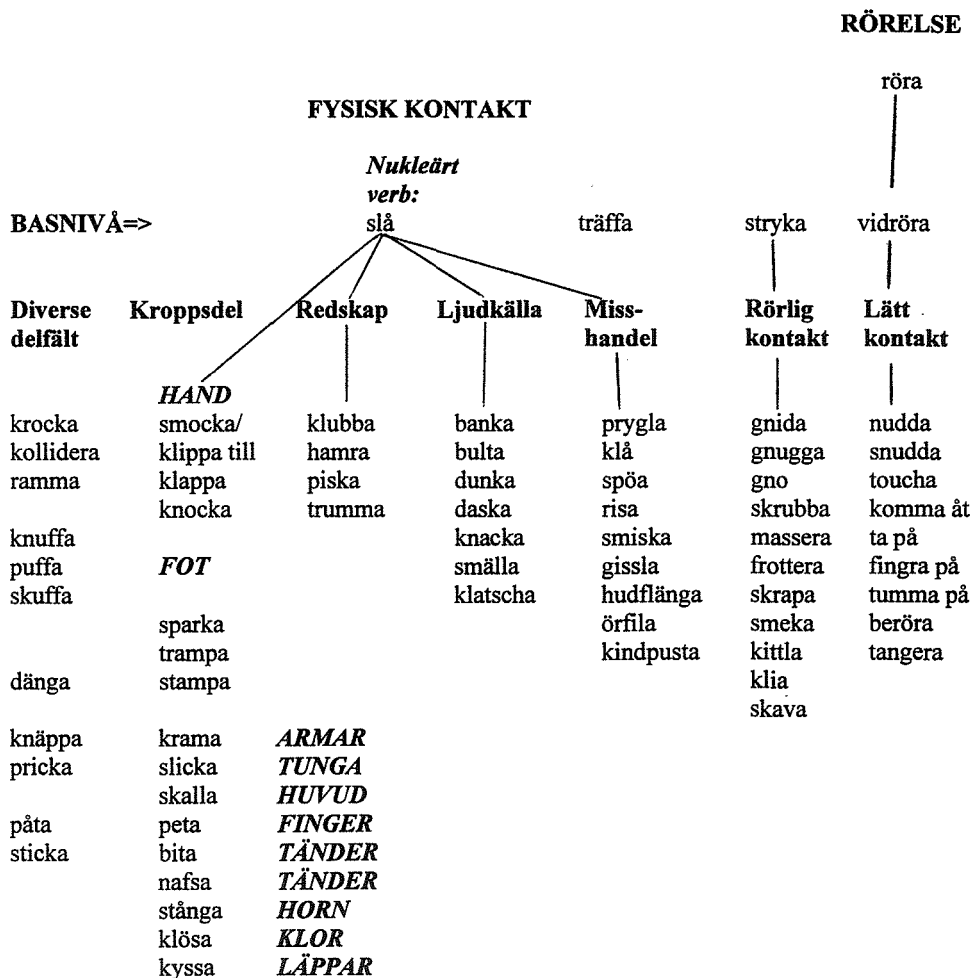
Per *strök* Pål ur rollistan.

Finns det några generella mönster för vilka betydelser verben inom ett fält såsom fysisk kontakt kan anta? På vad sätt bidrar den syntaktiska konstruktionen och betydelsen hos de ord som fyller olika syntaktiska positioner runt verbet till att en viss tolkning uppstår? Detta utgör uppsatsens frågeställning i ett nötskal.

2. En översikt av fältet Fysisk kontakt

I **Figur 1** visas en översikt av de fysiska kontaktverben uppdelade i semantiskt motiverade grupper.

Figur 1. Fysiska kontaktverb i svenskan. En översiktlig taxonomi.



Fältet är liksom övriga verbala semantiska fält organiserat kring ett nukleärt verb, verbet *slå*. Det nukleära verbet är semantiskt centralt genom att det har många hyponymer och förekommer i de flesta av de polysemimönster som är karakteristiska för fältet. Andra verb inom fältet med en mer specialiserad betydelse uppträder typiskt endast i en delmängd av de polysemimönster som är kännetecknande för det nukleära verbet. Verbet *slå* står också i en klass för sig inom fältet med avseende på textfrekvensen, vilket tyder på att det är lagrat i hjärnan på ett sätt som gör det mer tillgängligt vid sökning i det mentala lexikonet. Detta motiverar att *slå* ägnas särskild uppmärksamhet i analysen och att övriga verb inom fältet analyseras i relation till fältets nukleära verb.

Hyponymerna till *slå* specificerar den kroppsdel (*knocka* - KNYTNÄVE, *klappa* - HANDFLATA) eller det redskap (*hamra*, *piska*) kontakten uppkommer med. Bland

de verb som specificerar kroppsdelens är det främst de som förekommer under HAND som är hyponymer till *slå*. En särskild grupp utgörs av de verb som anger uppkomsten av ett ljud genom slag (*Per dunkade näven i bordet*) och som samtidigt kan fungera som rena ljudkälleverb (*det dunkar*). Särskilt rik är differentieringen mellan verb som beskriver misshandel (eller, framförallt i äldre tid, bestraffning) i olika former (*klå, spöa, prygla*). En särskild grupp fysiska kontaktverb utgörs av *stryka* och dess hyponymer, vilka beskriver hur något förflyttas längs med något samtidigt som det råder kontakt. (Utförligare om denna grupp i Viberg 1984.) De fysiska kontaktverben är nära besläktade med rörelseverben. Det mest generella av dessa i svenskan *röra* kan också fungera som fysiskt kontaktverb (*Rör inte glasskärvorna!*). Den prefixerade formen *vidröra* bildar tillsammans med *nudda* och *snudda* ett delfält som anger lätt beröring.

3. Den prototypiska betydelsen hos *slå*

I det här avsnittet kommer den lexikala representationen för *slå* i dess prototypiska betydelse att kortfattat beskrivas. Analysen sammanfattas i **Figur 2**.

Figur 2. Lexikal representation för den prototypiska användningen av verbet *slå*.

Syntaktisk ram:

NPZ ___ NP^W (prep NPY) (med NP^X)

Exempel: Per(z) slog Pål (w) (i ansiktet^y) (med knytnäven^x/en käpp^x)

Semantisk representation:

omarkerade värden:

HUMAN(z); HUMAN(w)

HAND(x) or HOLD(z,x)

BODYPART(y,w)

FAST(iii)

GREAT (d₁)

y utgör en kroppsdel hos w
förflyttningen i (iii) är snabb
kraften (d₁) är stor

(i) ACT(z,S)

z utför en handling S

(ii) STRETCH(ARM)

armen rätas ut

(iii) TRAVEL(x)

x förflyttas

(iv) HAPPEN(CONTACT(x,y))

det uppstår kontakt mellan x och y

(v) FORCE(x, y, d₁)

en kraft med styrkan d₁ riktas mot y

(vi) INTEND(z, DEFEAT or HURT(z, w))

z avser att besegra eller göra illa w

Ett verb sammanfattar ofta semantiskt en serie handlingar. Att öppna en dörr kan enligt ett utslitet mönster i en deckarroman beskrivas som att någon fattar tag i handtaget, vrider om detta och skjuter upp dörren. Det är en öppen fråga hur långt beskrivningen ska drivas. En förutsättning för att fatta tag i handtaget är t ex att armen sträcks ut och förs

mot handtaget. Då det gäller verbet *öppna* är det tveksamt om de typiska kroppsrörelserna då man öppnar en dörr ska specificeras i den semantiska representationen. Man kan ju t ex öppna en dörr genom att trycka på en knapp. Då man öppnar ett brev eller en konserverburk är kroppsrörelserna helt annorlunda. Verbet måste sägas vara neutralt med avseende på metoden. I satsen med fysiska kontaktverb som *Per slog in dörren* eller *Per sparkade upp dörren* fokuseras däremot metoden. En betydande andel av de fysiska kontaktverben kontrasterar med avseende på vilken kroppsdel som är involverad och den typiska rörelsen. Verbet *sparka* involverar t ex foten. Samtidigt kontrasterar det med verb som *trampa* och *stampa* med avseende på fotens rörelsemönster.

I typexemplet *Per slog Pål i magen* anger verbet *slå* att Per (z) utför en handling: ACT(z,S). Denna består i att armen sträcks ut: STRETCH(ARM). Att detta moment särskilt urskiljs motiveras av att *slå* kan användas som ett renodlat kroppsrörelseverb: *Per slog ut med armarna*. Att armen sträcks leder i sin tur till att handen eller något som hålls i handen (x) sätts i rörelse: TRAVEL(x) och kommer i kontakt med Påls mage (y): HAPPEN(CONTACT(x,y)). Kontakten i sin tur leder till att en kraft med en bestämd styrka (d_1) riktas från handen (x) mot magen (y): FORCE(x,y, d_1). Då objektet är konkret i största allmänhet (+/-human) kan kraften ge upphov till ett antal fysiska resultat, som kommer att behandlas litet utförligare längre fram: *Per slog omkull/ taktén på/ sönder trumman*. Då subjektet är humant tillkommer vid *slå* som vid många av de andra fysiska kontaktverben en överlagrad komponent som anger intentionen bakom handlingen. Kroppskontakt har ofta en kommunikativ innebörd (dunka i ryggen, sparka på smalbenen) eller uttrycker mer allmänt subjektets attityd gentemot objektet (agressivitet, ömhet). Verbet *slå* uttrycker i allmänhet att subjektet vill besegra eller göra illa objektet. Den förra tolkningen renodlas i exempel som *Per slog Pål i schack*, medan den senare är starkt implicerad då barn kommer till sina föräldrar och säger att någon slagit dem. Detta gäller i de mest typiska fallen. Specificerandet av en viss kroppsdel kan implicera andra innebörder av detta slag. Uttrycket *slå armarna om någon* uttrycker snarare ömhet eller vänskap. Humana objekt intar en särställning såtillvida att andra mål för ett slag inte gärna uppträder som direkt objekt. Små barn personifierar närmast då de säger saker som: *Jag ska slå det dumma bordet* (efter att ha gjort sig illa på det).

4. Subjektets roll för flertydigheten

Flertydigheten hos ett verb är till stor del en funktion av de semantiska egenskaperna hos de led som uppträder i den syntaktiska ramen, dvs. subjektet, objektet och övriga komplement. I det här avsnittet ska vi se hur subjektet påverkar den semantiska tolkningen. Ett humant subjekt bidrar ofta starkt till att ett exempel uppfattas som prototypiskt. Jämför t ex vid förflytningsverb som *gå*: *Flickan/ Bussen/ Klockan/ Tiden går*. Vid verben uppstår som här normalt ett kontinuum. Ett subjekt som *buss* saknar draget human/agent men undergår en konkret förflyttning, vid *klocka* rör det sig fortfarande om en konkret process, medan det vid *tid* som subjekt helt klart rör sig om en abstrakt process.

Beskrivningen av polysemin hos de fysiska kontaktverben grundar sig på en genomgång av samtliga exempel av *slå* och vissa andra av verben inom fältet i Språkbankens konkordanser baserade på svenska romaner 1980/81. Ett studium av vilka typer av substantiv som tenderar att uppträda som subjekt med särskilt hög frekvens avslöjar en hel del om ett verbs semantik. För ett nukleärt verb som *slå* vilket kan kombineras med ett vitt spektrum av substantiv är valet av subjekt en viktig faktor vid beskrivandet av verbets polysemimönster. För verb med en mer specialiserad betydelse gäller att vissa substantiv kan uppfattas som typiska subjekt som bidrar till karakteristiken av verbets betydelse. Verbet *träffa* har t ex ofta ett subjekt som betecknar en projektil (kula, pil, granat, torped) medan fordon av skilda slag är utmärkande för *krocka*. Iakttagelserna kan sammanfattas i den skala över avtagande prototypikalitet hos subjektet som presenteras i **Figur 3**.

Figur 3. Avtagande prototypikalitet hos representativa klasser av subjekt till Fysiska kontaktverb

| HUMAN => | INSTRUMENT => | NATURKRAFT => | PERCEPTUELLA => FENOMEN | MENTALA FENOMEN |
|----------|---|-------------------------|----------------------------|--------------------|
| | KLOCKA FORDON PROJEKTIL (KROPPSDEL) (REDSKAP) | VIND VATTEN (ELD) | LJUS LJUD LUKT | TANKE, KÄNSLA |

BOKSTAVLIG

- 1) *Intentionell*
- 2) *Självständig förflyttning*
- 3) *Konkret förnimbar med flera sinnen*

METAFORISK

Rent semantiskt är avvikelserna från det prototypiska fallet rätt liten, då subjektet utgörs av en kroppsdel:

I det djupa mörkret strax efter midnatt slog en hand varligt på Adams sovrumsfönster.

Syntaktiskt avviker exemplet från det grundläggande mönstret genom att instrumentet *hand* gjorts till subjekt i stället för en agent som inte explicit markeras i satsen. Även om subjektet *hand* inte kan tillskrivas en intention, så underförstås ändå att det fanns någon som slog med sin hand mot fönstret, förmodligen för att dra till sig uppmärksamheten. Exempel med en kroppsdel som subjekt har låg frekvens men tycks användas just i situationer där det inte är klart vem som utför handlingen.

Mer påtaglig för verbet *slå*:s del är den stegvisa förskjutning av betydelsen som sker vid följande typer av subjekt:

Naturkraft

Han öppnade dörren till styrhytten och vinden slog emot honom med våldsamt kraft
Regnet slog mot fönstren.

Perceptuella fenomen

Håkans snarkningar slog emot mig i trappan.

LJUD

En stark, unken lukt av mögel slog emot mig när jag steg in. LUKT

Mental Metaforisk princip 1: En idé är en kraft

Plötsligt slog en tanke ner i Ann.

Nej, kanske var hon inte så ful, medger modern.

Det slog mig också när jag tittade igenom albumet.

Hans idéer slog inte igenom.

Då subjektet utgörs av en naturkraft (vind, regn, våg) rör det sig fortfarande om en rent fysisk kontakt. Skillnaden gentemot det prototypiska fallet är främst att subjektet inte kan tolkas som en agent som utför en intentionell handling. Satsers vars subjekt utgörs av perceptuella fenomen som ljus, ljud och lukt uppfattas som något mer abstrakta, vilket kan låta paradoxalt med tanke på att konkret tidigare förklarades vara något som kunde förnimmas direkt med våra sinnen. Ett slags kontakt uppstår även då t ex ett starkt ljus plötsligt riktas mot ansiktet (*det starka skenet slog emot mig*). Men kontakten uppfattas inte som i det prototypiska fallet med känselsinnet (hos den individ som fungerar som Erfarare). Till skillnad från de mer konkreta fallen rör det sig om fenomen som bara kan upplevas med ett sinne åt gången. Kraften i kontakten är dessutom mycket begränsad. Perceptuella fenomen kan inte som en naturkraft flytta föremål (*vinden slog omkull mig*) eller orsaka mer påtagliga fysiska förändringar (t ex *slå sönder*).

Då subjektet utgörs av ett mentalt fenomen kan man inte längre tala om fysisk kontakt. Sådana exempel kräver för sin tolkning närmast en metaforisk princip av det slag som beskrivs i Lakoff & Johnson (1980): "En idé är en kraft". Då man säger att man slås av en tanke anger det att tanken kommer plötsligt. Vanligtvis anger *slå* också att tanken väcker en fysiologisk reaktion. Semantiskt sett innebär detta att även övergången från perceptuella till mentala fenomen i subjektshierarkin bildar ett kontinuum och inte är helt abrupt.

Metaforen måste anses vara helt konventionaliserad i de ganska många exemplen av typen *det slog mig att* + S, *tanken slog henne att* + S. Det förekommer dock att metaforen görs mer levande genom en direkt anknytning till konkreta typer av fysisk kontakt i form av en liknelse (av typen: som en/ett N):

Varför? Det var den första tanken som slog mig
-- slog mig som ett knytnävsslag mellan ögonen --

Deras ovilja slog emot honom som en kall fläkt,

Analysen av subjektivalet vid verbet *slå* kan kort jämföras med motsvarande mönster hos ett fysiskt kontaktverb med närstående betydelse, nämligen *träffa*, vilket här kommer att behandlas endast då det fungerar som fysiskt kontaktverb. Den mest frekventa betydelsen hos detta verb hör till ett annat fält nämligen Social kontakt (*Per träffade Pål*

Åke Viberg

på partyt). Subjektet är i denna användning alltid humant. Som ett fysiskt kontaktverb fokuserar *träffa* själva resultatet och har närmast den semantiska representationen:

(i) HAPPEN(CONTACT(x,y))

Verbet kan även som ett fysiskt kontaktverb ha ett **Humant subjekt**:

Lennart slog ut mot Edi en gång till. Han träffade honom på kinden med en fullt hörbar smäll.

Det mest typiska subjektet utgörs emellertid av substantiv som *kula, pil, granat, bomb, torped* vilka sammanfattande kan betecknas som **Projektiler**:

Två gånger sköt han och båda pilarna träffade sitt mål.

Jag fick ett rekordkast och spjutet träffade den elektriska ledningen till rävfarmen.

Även andra substantiv än de som lexikalt betecknar en projektil kan förekomma i en liknande funktion:

Uglik tog en klump snö och kastade den så att den träffade björnen i baken.

Precis som vid *slå* bildar perceptuella fenomen som subjekt semantiskt en övergång mot rent abstrakta bildliga användningar:

Perceptuella fenomen

I sista sekund sköt en ljusstråle fram och träffade mig mitt i ansiktet.

De gick in på Sputnikbaren för en stund. Musiken träffade dem som en stöt för huvudet.

Mental Metaforisk princip 2: "Ord är projektiler"

och Villons gyckel träffade honom på en ganska öm punkt

Metaforerna kan även vid detta verb göras mer levande genom explicita liknelser:

Hennes yttrande träffade honom som en stöt

Varje intryck träffade honom som ett skarpt hugg

5. Polysemi knuten till objektet och andra komplement

Objektet till verbet *slå* kan tillhöra semantiskt skilda klasser som på olika sätt interagerar med verbets grundbetydelse för att ge upphov till ett vitt spektrum av sekundära betydelser. Huvuddelen av dessa kan beskrivas som ett tillägg av en resultativ komponent.

Det resultat som kodas språkligt vid de fysiska kontaktverben följer lagarna för vad som kan kallas vardagslivets fysik. Vardagslivets fysik säger oss att vid uppkomsten av en fysisk kontakt bildas en kraft. En kraft kan i sin tur ge upphov till att objektet förflyttas (*Per slog iväg bollen med knytnäven*), går sönder (*Per slog sönder flaskan*) eller sänder ut ljud (*Per slog takten på trumman*). Vad som sker är att predikatsuttrycket i sin helhet övergår till ett annat semantiskt fält. I exemplet *Per slog iväg bollen* kan *slå* parafraseras 'förflytta med ett slag' och verbet kontrasterar med förflyttningsverb som *kasta*.

Fallet då objektet förflyttas kan illustreras med följande autentiska exempel i vilket variabler lagts till inom parentes:

han (z) gick i tankar och slog med käppen (x) en liten sten (y)
framför sig längs gatan.

Den situation som verbet *slå* beskriver i detta exempel består i att z rör sin arm och hand i vilken käppen hålls på ett sådant sätt att käppen kommer i kontakt med stenen som därvid förflyttas. Även om *slå* fungerar som ett förflyttningsverb, så föreligger fortfarande betydelsen fysisk kontakt. Förflyttningsskomponenten har adderats till verbets grundbetydelse. Den semantiska representationen visas i **Figur 4**. (Något förenklat. Bl a underförstås alla CAUSE utom det sista.)

Figur 4. Den semantiska representationen för *slå* som Förflyttningsverb

| | | |
|---------|----------------------------|---|
| (i) | ACT(z,S) | z utför en handling |
| (ii) | STRETCH(ARM) | armen sträcks |
| (iii) | TRAVEL(x) | x förflyttas |
| (iv) | HAPPEN(CONTACT(x,y)) | det uppstår kontakt mellan x och y |
| (v) | FORCE(x,y,d ₁) | en kraft med styrkan d ₁ riktas mot y från x |
| (vi) a) | CAUSE((v), TRAVEL(y)) | vilket leder till att y förflyttas |

Komponenten TRAVEL(y) som lagts till i (vi) utgör kärnkomponenten i fältet Förflyttning (Viberg 1992b). De övriga fallen kan beskrivas genom att den resultativa komponenten (vi) byts ut mot kärnpredikatet i andra fält. En översikt av ett antal sådana fall där *slå* antar olika sekundära betydelser visas i **Figur 5**.

Figur 5. Viktigare typer av sekundära betydelser hos slå

| | Semantiska komponenter | Semantiskt fält <i>Humant subjekt</i> | <i>Jämför med annat subjekt:</i> |
|---------|--|--|---|
| (i) | ACT(z,S) | Stationär rörelse: Kroppsrörelse | |
| (ii) | STRETCH (ARM) | Per slog ut med armarna. | Blommorna slog ut. |
| (iii) | TRAVEL(x) | | |
| (iv) | HAPPEN(CONTACT(x,y)) | Renodlad kontakt Per slog emot ett bord. | Vågorna slog mot stranden. |
| (v) | FORCE(x,y,d ₁) & CAUSE(d ₁ ,(vi)) u: GREAT (d ₁) | | |
| (vi) a) | TRAVEL(y) | Objektiv förflyttning Per slog bort bollen. | |
| (vi) b) | (FROM(TRAVEL))(SOUND,y) | Ljudkälla Per slog takten. | Det slog i dörrarna. Klockan slog fem. |
| (vi) c) | HAPPEN(notCONNECTED(y)) | Diskonnexion Per slog sönder flaskan. Per slog gräsmattan. | |
| (vi) d) | HAPPEN(FEEL(w,PAIN)) | Kroppsförnimmelse Per slog sig (i knät). | Pulsen/Hjärtat slår. |
| (vi) e) | HAPPEN(notALIVE(w)) | Organiskt liv Per slog ihjäl Pål. | Björnen slog ett får. |
| (vi) f) | HAPPEN(EXIST(FIRE)) | Produktion Per slog eld på en tändsticka. | |
| (vi) g) | DEFEAT (z,w) | Social interaktion Per slog Pål i schack. Napoleon slog preussarna. | |

Uttrymmet tillåter bara några korta kommentarer till ett par av fallen. Det finns i svenskan ett ganska stort antal renodlade **ljudkälleverb** av typen *susa*, *surra*, *mullra*. Syntaktiskt kännetecknas de av att de kan uppträda i ramen: det ____ prep NP^y, där NP:et betecknar källan varifrån ljudet kommer (*det susar i vassen*, *det surrar i kupan*). Semantiskt delar ljudkälleverben följande komponent:

- (i) (FROM(TRAVEL))(SOUND, y) ett ljud kommer från y

Denna komponent ingår också i den semantiska representationen för de hyponymer till *slå* som listats under Ljudkälla i figur 1. Flera av de övriga fysiska kontaktverben kan

också uppträda i den för ljudkälleverben karakteristiska syntaktiska ramen och har då adderat (i) till sin grundbetydelse: *det slår i dörrarna, det stampar i farstun*. Det finns också fall där objektet till *slå* anger att verbet ska tolkas som ett ljudkälleverb som i följande autentiska exempel:

Henry visslade fortfarande på *La Cucaracha* och slog takten mot skåpdörrarna i köket så det dånade genom hela huset.

Den semantiska representationen är som i figur 4 med undantag av att (vi) har formen:

(vi) (b) CAUSE((v), HAPPEN((FROM(TRAVEL))(SOUND, y)))

Ytterligare ett fall föreligger i ett exempel som det följande, där *slå* fungerar som ett **diskonnexionsverb**:

Där tog jag en glaskaraff och slog av dess hals över en stolsrygg. (R76)

Kontakten orsakar här att glaskaraffen går sönder vilket beskrivs med komponenten notCONNECTED(x) som utgör kärnpredikat i fältet Diskonnexion (Viberg 1985). Till detta fält hör verb som *skära, klippa, bryta, krossa*, vilka anger att ett föremål delas i bitar eller går sönder. Denna betydelse kan också som i exemplet ovan signaleras av partiklar som *av, sönder, itu*. Betydelsen signaleras dock inte explicit vid verbet *slå* i exempel som de följande:

Vilhelm slog dikeskanterna med lie, och jag körde slåttermaskinen. (R76)

Caesar Lund gick därute och slog gräset utan att titta upp (R76)

I det översta exemplet anges explicit att instrumentet är en lie, men som framgår av det undre exemplet kan den betydelsen underförstås då objektet utgörs av *gräs* och besläktade uttryck (*gräsmatta, dikesren, slänt* - platser där det växer gräs). Litet förenklat kan den syntaktiska ramen och den semantiska representationen återges som i **Figur 6**.

Figur 6. Den lexikala representationen för 'slå gräs'

Syntaktisk ram: NP^Z ___ NP^Y (med NP^X) Per slog gräset (med lie).

HUMAN(z), GRASS(y), SCYTHE(x)

- (i) ACT(z,S)
- (ii) TRAVEL(x)
- (iii) HAPPEN(CONTACT(x,y))
- (iv) CAUSE((iii), HAPPEN(notCONNECTED(y)))

Att instrumentet är lexikaliserat och kan underförstås i kombination med vissa bestämda typer av objekt har paralleller hos andra verb. Exemplet *Per slog gräset* är analogt med *Per strök taket* och *Per strök skjortan*, i vilka instrumenten PENSEL och STRYKJÄRN är underförstådda (se Viberg 1992a). Sådana specialbetydelser som

brukar finnas listade i en ordbok kan således beskrivas genom att addera vissa komponenter till den semantiska representation som svarar mot verbets grundbetydelse.

6. Slutord

Frekventa och centrala verb som *slå* har en komplex lexikal representation som täcker ett stort antal skilda betydelser. Det har i den här uppsatsen bara varit möjligt att ge en översiktlig beskrivning av några av de polysemimönster som är karakteristiska för *slå* och de andra fysiska kontaktverben. Vad som sagts antyder dock hur pass komplicerad en uttömmande beskrivning måste bli. Samtidigt gäller att flertydigheten följer systematiska mönster. Även om de sekundära betydelserna ofta inte är helt förutsägbara och restlöst kan härledas ur ett litet antal primitiver och principer för hur dessa kombineras, så förenklas en explicit beskrivning av betydelsen i väsentlig grad av att många betydelser som vid första påseendet kan verka orelaterade vid ett närmare betraktande kan beskrivas genom en lätt modifikation av den grundläggande semantiska representationen. Om vi i enlighet med en numera vanlig uppfattning tänker oss att lexikonet i sin helhet är realiserat som ett nätverk, sker en väsentlig besparing genom att bara några länkar är unika för de enskilda betydelserna medan en betydande del av det semantiska nätverket är gemensamt för olika betydelser av ett verb som *slå*. Denna centrala del av nätverket delas också i stor utsträckning av andra verb som hör till fältet.

Not

Arbetet med denna uppsats har möjliggjorts genom generöst stöd från Humanistisk-samhällsvetenskapliga forskningsrådet (HSFR) i form av en tjänst i lexikalisk forskning.

Litteratur

Autentiska exempel har hämtats ur: *Språkbanken*, Institutionen för svenska språket, Göteborgs universitet. Huvudsakligen har *Romaner 1980-81* utnyttjats. I några få fall, då *Romaner 1976* utnyttjats följs exemplen av källangivelsen (R76).

Lakoff, G & Johnson, M (1980), *Metaphors we live by*. University of Chicago Press, Chicago.

Miller, G A & Johnson-Laird, P N (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.

--- (1984). Fysiska kontaktverb i svenskan. En skiss. I: *Svenskans beskrivning 14*. Lund.

--- (1985). Hel och trasig. En skiss av några verbala semantiska fält i svenskan. I: *Svenskans beskrivning 15*. Göteborgs universitet.

---(1992a). Tvärspråklig lexikologi med svenskan i centrum. I: Fjeld, R (red.), *Nordiske studier i leksikografi. Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1*. Oslo.

---(1992b). Universellt och språkspecifikt i svenskans ordförråd. *Tijdschrift voor Scandinavistiek. Nr 2, 1992*.

Oversigt over forfatterne

- Olaf Almenningen* Institutt for nordistikk og litteraturvitenskap, Avdeling for leksikografi, Postboks 1001 Blindern, N-9315Oslo.
- Jens Axelsen* Gyldendal, Klareboderne 3, DK-1001 København K.
- Anna Braasch* Center for Sprogteknologi, Københavns Universitet, Njalsgade 80, DK-2300 København S.
- Anna-Lena Bucher* Tekniska nomenklaturcentralen, Västra vägen 9 C, S-171 46 Solna.
- Ulla Clausén* Svenska språknämnden, Lundagatan 42. uppg. 5, S-117 27 Stockholm.
- Anne Duekilde* ODS-Supplementet, Det Danske Sprog- og Litteraturselskab, Københavns Universitet, Njalsgade 80, DK-2300 København S.
- Dorthe Duncker og Ole Norling-Christensen* Den Danske Ordbog, Det Danske Sprog- og Litteraturselskab, Københavns Universitet, Njalsgade 80, DK-2300 København S.
- Sv. Eegholm-Pedersen* ODS-Supplementet, Det Danske Sprog- og Litteraturselskab, Københavns Universitet, Njalsgade 80, DK-2300 København S.
- Ruth Varvedt Fjeld* Institutt for nordistikk og litteraturvitenskap, Avdeling for nordisk språk og litteratur, Postboks 1013 Blindern, N-0315 Oslo.
- Anders Geertsen* Munksgaards Forlag, Nørre Søgade 35, DK-1370 København K.
- Maria Toporowska Gronostaj* Inst. för svenska språket, Göteborgs universitet, S-412 98 Göteborg.
- Ebba Hjorth og Knud Troels Thomsen* Den Danske Ordbog, Det Danske Sprog- og Litteraturselskab, Københavns Universitet, Njalsgade 80, DK-2300 København S.
- Karl Hyltdgaard-Jensen* Lillkullegatan 14, S-412 74 Göteborg.
- Claus Bo Jørgensen* TEXTware A/S, Hørsholmsgade 20, DK-2200 København N.
- Kjeld Kristensen* Den Danske Ordbog, Det Danske Sprog- og Litteraturselskab, Københavns Universitet, Njalsgade 80, DK-2300 København S.
- Hans-Peder Kromann* Handelshøjskolen i København, Dalgas Have 15, DK-2000 Frederiksberg.
- Guðrún Kvaran* Orðabók Háskólans, Neshaga 16, IS-107 Reykjavík.
- Birgitta Lindgren* Svenska språknämnden, Lundagatan 42, uppg. 5, S-117 27 Stockholm.
- Carl-Erik Lundblad* Svenska Akademiens ordbok, Box 3, S-221 00 Lund.
- Erika Lyly* Svenska Språknämnden, Lundagatan 42, uppg. 5, S-117 27 Stockholm.
- Sven-Göran Malmgren* Göteborgs universitet, Inst. för svenska språket, S-412 09 Göteborg.
- Nina Martola* Institutionen för de inhemska språken, Sörnäs strandväg 25, SF-00500 Helsingfors.

Oversigt over forfatterne

Hans Kristian Mikkelsen Handelshøjskolen i Århus, Russisk institut, Fuglesangs Allé 4,
DK-8210 Århus V.

Sandro Nielsen Handelshøjskole Syd, Østervang 2, DK-6800 Varde.

Kristina Nikula Puolalaparken 1 b B 12, SF-20100 Åbo.

Kerstin Norén Göteborgs universitet, Inst. för svenska språket, S-412 98 Göteborg.

Anders Drejer Nygaard TEXTware A/S, Hørsholmsgade 20, 2200 København N.

Christian-Emil Ore Universitetet i Oslo, Dokumentasjonsprosjektet, Boks 1102 Blindern, N-0317 Oslo.

Karen Margrethe Pedersen Institut for dansk dialektforskning, Københavns Universitet,
Njalsgade 80, DK-2300 København S.

Inger Hesselin Rider Norstedts Förlag AB, Norra Bankogränd 2, Box 2052, S-103 12
Stockholm.

Bo Svensén Svenska Akademien, Box 2118, S-103 13 Stockholm.

Åke Viberg Stockholms universitet, Centrum för tvåspråkighetsforskning, S-106 91
Stockholm.

Ordbogsregister

(som anført af forfatterne – dansk alfabetisering)

A

Árni Bödvarsson. 1963. Íslensk orðabók handa skólum og almenningi. Reykjavík. (Anden udgave forøget, Reykjavík 1983) 157

Ásgeir Blöndal Magnússon. 1989. Íslensk orðsifjabók. Reykjavík 157

B

Björn Halldórsson. 1814. Lexicon Islandico-Latino-Danicum Björnnonis Haldorsonii. Vol. I-II. Havniæ 157

Björn Halldórsson. 1992. Orðabók. Íslensk-latenesk-dönsk. Eftir handriti í Stofnum Árna Magnússonar í Kaupmannahöfn. Ny útgáfa. Jón Aðalsteinn Jónsson sá um útgáfuna. Reykjavík 157

Blinkenberg, Andreas og Poul Høybye: Dansk-fransk ordbog / Dictionnaire Danois-Français. 4. udg. ved Jens Rasmussen & al. Bind 1-2. København 1991 68

BS 3558 (1980) Glossary of rubber terms. London, British Standards Institution 277

C

Clausens tekniske ordbøger. Engelsk-dansk 37

Collins COBUILD English Language Dictionary. London & Glasgow 1987 37, 107, 122, 171.

Crystal, David 1991: A Dictionary of Linguistics and Phonetics, 3rd ed., Blackwell, Oxford 203

D

Dansk-norsk ordliste. Bokklubben Norden och Nordiska språksekretariatet. Oslo 1989 165

Dansk-svensk ordlista. Bokklubben Norden och Nordiska språksekretariatet. Oslo 1988 165

Dansk-Tysk Ordbog. Udarbejdet af J.J. Siegler. 431 s. Berlingske Forlag, København 151

Dansk-Tysk Ordbog. Undervisning. Af Grethe Hjorth. Sproglig konsulent Ingeborg Zint. Hovedredaktør Jens Axelsen. 1. udgave. Gyldendal, København 151

Den Danske Ordbog (DDO) 1-6 (under udarbejdelse) 122, 138

Deutsches Rechtswörterbuch. Hg. v. der Preussischen Akademie der Wissenschaften. Weimar 1914 ff., jetzt von der Heidelberger Akademie der Wissenschaften 128

Deutsches Universalwörterbuch. 2., völlig neu bearb. u. stark erw. Aufl. Hrg. u. bearb.v. Wiss. Rat u. d. Mitarb. der Dudenred. unter Leitung von Günther Drosdowski. Dudenverlag. Mannheim, Leipzig, Wien, Zürich. 1989 122, 151

dtv-Wörterbuch der deutschen Sprache. Hrsg. v. Gerhard Wahrig in Zus. m. zahlr. Wiss. u. a. Fachl. Deutscher Taschenbuchverlag, München 1978 151

E

Egil Pettersen: Bergensordboken. Bergen 1991 16

Eiríkur Jónsson. 1863. Oldnordisk Ordbog ved det kongelige nordiske Oldskrift-Selskab. Kjöbenhavn 157

Ensk-føroysk orðabók 240

F

Fritzner, Johan. 1867. Ordbog over det gamle norske Sprog. Kristiania 157

Frühneuhochdeutsches Wörterbuch. 1986. Hg. v. Robert R. Anderson, Ulrich Goebel, Oskar Reichmann. Berlin, New York 128

G

Gerhard Wahrig: Das große deutsche Wörterbuch. Hrsg. in Zus. m. zahlr. Wiss. u.a. Fachl. Mit einem "Lexikon der deutschen Sprachlehre". Bertelsmann, Gütersloh 1966 151

Grimm: Deutsches Wörterbuch 69

Guðbrandur Vigfússon. 1874-76. An Icelandic-English Dictionary based on the MS. Collections of the late Richard Cleasby enlarged and completed by Gudbrand Vigfusson. With an Introduction and Life of Richard Cleasby and George Webbe Dantsen. Oxford. (2nd Ed 157

Guðmundur Andrússon. 1683. Lexicon Islandicum Sive Gothicae Runæ vel Lingvæ Septentrionalis Dictionarium ... Havnæ 158

Gunnlaugur Oddsson. 1819. Orðabók, sem inniheldr flest fágæt, framandi og vand-skilinn orð, er verða fyrir í dönskum bókum. Kaupmannahöfn 158

Gunnlaugur Oddsson. 1991. Orðabók, sem inniheldur flest fágæt, framandi og vand-skilin orð, er verða fyrir í dönskum bókum. Ny útgáfa með íslenskri orðaskrá. Jón Hilmar Jónsson sá um útgáfuna ásamt Þórdísi Úlfarsdóttur. Reykjavík 158

H

Handwörterbuch zur deutschen Rechtsgeschichte. Hg. v. Adalbert Erler und Eckehard Kaufmann. Berlin 1971f 128

Henriksen, T. (1991): Juridisk ordbog spansk-dansk. København, Gad 278

Hrefna Arnalds, Ingibjörg Johannesen (red.) 1992. Dönsk-íslensk orðabók. Reykjavík 158

I

Illustrerad svensk ordbok. 1956. Stockholm 184

J

Jørgensen, Merete K., Nielsen, Marita Akhøj, Sørensen, John Kousgård
Gammeldansk Ordbog. Det danske Sprog- og Litteraturselskab, Præsentationshæfte
21. København 1992 128
Jysk Ordbog 248

K

Kaufmann, U. & Bergenholtz, H. (1992): Genteknologisk ordbog. Dansk-engelsk / engelsk-dansk molekylærbiologi og DNA-teknologi. København, Gad 278
Kjærulff-Nielsen, B.: Engelsk-Dansk Ordbog. Gyldendal, København, 1985 37
Konráð Gíslason. 1851. Dönsk orðabók með íslenzkum þýðingum. Kaupmannahöfn
158

L

Longman Dictionary of Contemporary English. (1987) Harlow, Essex 107, 171, 236236

M

Mittelniederdeutsches Handwörterbuch. Hg. v. Agathe Lasch, Conrad Borchling und
Gerhard Cordes, jetzt v. Dieter Möhn. Neumünster 1928 ff 129
Molbech, Chr.: Dansk Ordbog, 1833 18

N

Nationalencyklopedins ordbok. Under utarbetande. Bra Böcker, Höganäs 171
Nielsen, Feodor: Anglo/amerikansk-dansk retsordbog. København,
Nyt Nordisk Forlag Arnold Busk 1975 216
Nielsen, Sandro: Engelsk-dansk juridisk basisordbog. København, Munksgaard 1993
216
Nordisk förvaltningsordbok. Rev. uppl. Nordiska rådet. 1991 165
Nordisk leksikografisk ordbog, NLO 193
Nordisk leksikografisk ordbok (under udarbejdelse) 78
Norsk-dansk ordliste. Bokklubben Norden och Nordiska språksekretariatet. Oslo 1989
165
Norsk-svensk ordlista. Bokklubben Norden och Nordiska språksekretariatet. Oslo 1988
165
Norstedts dansk-svenska ordbok. Utarbetad av Bertil Molde under medverkan av Niels
Ferlov. 3 uppl. Stockholm, Norstedts 1980 165
Norstedts norsk-svenska ordbok. Utarbetad av Birgitta Fowler och Kjell Ivar Vannebo
i samarbete med Lena Moberg vid Svenska språknämnden. Stockholm, Norstedts
1992 165

O

ODS, Ordbog over det Danske Sprog, 1-28, København 1918-56: 53, 68, 69, 122, 248

ODS-Supplementet, bind 1,A(Aa)-Bh. 1992 69
Ømålsordbogen 248
Ordbok över Svenska Språket (SAOB) utgiven av Svenska Akademien, Band 1ff. Lund
1898ff 53, 69, 122
Oxford Advanced Learner's Dictionary of Current English, Encyclopedic Edition. Ox-
ford 1992 122
Oxford English Dictionary (OED) 69

P

Politikens Nudansk Ordbog. Redaktion Christian Becker-Christensen, Peter Widell.
14. udgave. København 1990 (11953) 151

R

Ragnhild Paulsen: Ordbok over nøttlandsmålet omkring 1900. Tønsberg 1981 16
Repp & Ferrall: Dansk-Engelsk Ordbog, 1848 18
Retskrivningsordbogen. Udgivet af Dansk Sprognævn. København 1986 68

S

Samnordisk ordbog/ordbok. Red. Andreas Fynning. 2 uppl. København
Nyt Nordisk Forlag 1971 165
SAOL=Svenska Akademiens ordlista. 10, 11. uppl. 1973, 1986 Stockholm 184
Schiller, K. und Lübber A.: Mittelniederdeutsches Wörterbuch. Bd. 1-6. Bremen 1875-
81 129
Skibsted, S.: Teknisk engelsk-dansk ordbog. København, 1971 37
Skjerk, Ebbe: Bilteknisk ordbog. Engelsk-Dansk. Teknisk Forlag, 1991 37
SMS 508/TNC 78 (1985). Verkstadsteknisk ordlista. Skärande bearbetning. Stock-
holm, Sveriges Mekanstandardisering & Tekniska nomenklaturcentralen 278
SS 01 16 01/TNC 82 (1984). Dataordboken. Utgåva 3. Stockholm
SIS - Standardiseringskommissionen i Sverige 278
Svensk handordbok, 1966 47
Svensk konstruktionsordbok (SKO) (under udarbejdelse) 47, 173
Svensk ordbok (1986), Språkdata. Göteborgs universitet. Esselte Studium 107, 172,
226, 236
Svensk ordbok. 1948. Svenska dövstumläraresällskapet, Lund 172
Svensk-dansk ordliste. Bokklubben Norden och Nordiska språksekretariatet. Oslo 1989
165
Svensk-norsk ordliste. Bokklubben Norden och Nordiska språksekretariatet. Oslo 1989
165
Sören Sörensson. 1984. Ensk-íslensk orðabók með alfræðilegu ívafi. Reykjavík 158

T

Teknisk ordbog. Engelsk-dansk. L&H Ordbøger. København, 1991 37

The American Heritage Dictionary of the English Language. Third Edition. Boston 1992 122

Thomas Bloch: Glossemata Tellemarchia. Fyresdal 1698. Oslo 1956 16

TNC 81 (1984). Energiordlista. Stockholm, Tekniska nomenklaturcentralen 278

TNC 91 (1990). Svensk-engelsk byggordbok. Solna, Tekniska nomenklaturcentralen 278

TNC 93 (1992). Transportordlista. Solna. SIS-STG & Tekniska nomenklaturcentralen 278

Tysk-dansk, dansk-tysk. Redaktion Henning Poulsen. Konsulenter Hans- Peder Kromann, Monika Jensen. 1. udgave. Gad, København 1987 151

V

Vinterberg, Hermann og C.A. Bodelsen: Dansk-engelsk Ordbog. 3. udg. ved Viggo Hjørnager Pedersen. København 1990 68, 122

W

Woordenboek der Nederlandsche Taal 69

Wörterbuch der deutschen Gegenwartssprache. Herausgegeben von Ruth Klappenbach und Wolfgang Steinitz. Band 1-6. Berlin 1964-1977 122

Wörterbuch der mittelhochdeutschen Urkundensprache auf der Grundlage der altdeutschen Originalurkunden bis zum Jahr 1300. Unter Leitung von Bettina Kirschstein und Ursula Schulze erarbeitet von Sibylle Ohly und Peter Schmitt. Berlin 1986f 129

Stikordsregister

(dansk alfabetisering)

A

afledninger 67
akronymmer 83
allmennord 81
allmännospråkliga uttryck 43
alordbok 275
alordbog 205
alordbogskonseptet 216
analysestruktur 67
antonym 170
användning 168
artikelskelett 68,109
atomar artikel 237
attribut 62
automatisk lemmatisering 183
avledning 167

B

begrebsrelasjoner 36
begrepp 43, 44
begreppsharmonisering 46, 264
begreppsinventering 264
begreppsseleksjon 264
begreppssystematik 263
begreppssystematisering 264
betydelseanalyse 101
betydelsebeskrivelse 168
betydningsdifferentiering 113, 114
betydningsomfang 34
betydningsopdeling 109, 110, 111
bildlig 169, 170
brugsbestemmelser 75
bøjningsform 167
bøjningsmønster 65
bøjningsoplysninger 64
bøjningsparadigmer 67

brugsbestemmelser 75

C

COMPULEXIS 87
CX-Basic 91
CX-Master 91

D

Dansk Sprognævn 71
datering 53
DDO 61
deduktiv 196
definition 46, 168
delekvivalent 80
delfagsordbøger 216
Den Danske Ordbog 61
deskriptiv ordliste 41, 44
dialektordbøger 248
DTD 95
dubbelt sammansatta ord 179
dubletter 81
döva 166

E

EES-avtalet 38
eftertext 265
egentlig ækvivalent 208
EGs regelverk 38, 45
eksempler 144
ekvivalens 78, 101
ekvivalensrelasjon 78
elektroniske ordbøger 130
emneklassifisering 109
encyklopædi 205
encyklopædiske opplysninger 248
encyklopedisk 170

encyklopedisk systematik 263, 271, 275
entitet 62
entitet-relationsdiagram 62
episk ækvivalent 208
ettordsenheter 84
etymologisk datering 57
Excerptering 39
explicit systematik 263, 265, 276

F

fackspecifikt begrepp 43
fackspråkliga idiom 41
fagintroduktion 214
fagordbog 204
fagtermer 83
falske lemmaer 82
fasta fraser 48, 173
feltopdelt data 87
feltstruktur 91
fleksiver 67
flerordsforbindelser 209
Flerordstermer 43
flertydige konstruktioner 133
flertydighed 31
flertydighet 283
folkelig terminologi 251
forklarende ækvivalent 208
fornsvenskan 125
fras 41
fraseangivelse 209
fraser 39, 43
fremdatering 56
fremmedord 73
frilansare 258
funktioner 194
fysiska kontaktverb 279
fältet Diskonnexion 289
förflyttningsverb 287
förled 179
förstebelæg 58
förtext 265

G

gammeldansk 124
genbrug 108
GestorLEX 130
gradbøjning 67
grammatik 143
grundbetydelse 169

H

header 140
heterograf 63
hierarkisk ordning 174
hierarkisk struktur 131
homografi 183
homografseparation 67
hovedfagsordbøger 216
huvudoppslagsord 167
hybridord 80
hänvisningar 270

I

idiom 48, 169
illustration 167
implicit systematik 263, 264, 276
induktiv 196
internasjonalismer 79
introduktion till ämnesområdet 265
invandrare 170

J

juridisk begreppssystem 125
juristlingvist 45
jämförelser 49

K

klassifikation 200
klassningskod 270
kollokationer 33, 39, 41, 43, 73, 173
komparativ 67
kompetence 66
konkordansprogrammel 67
konkordansregister 40

konsensus-database 91
konstruktion 144
konstruktioner 144
kontextuella översättningar 41
kontrastiv 227
korpus 61, 139, 183
korpusbaserad ordlista 41
korpusbaseret ordbogsarbejde 35
Kreativ variation 52
krydshenvisninger 208, 251
krydshenvisningsartikler 215

L

leksikografiens fagsprog 194
leksikografiske behov 206
leksikografiske symboler 207
leksikonafsnit 250
lemma 81, 125, 205
lemmasektion 113
lemmasektion til fagordbøger 196
lemmatiseringssystem 183
lexfraser 175
lexikal variation 51
lexikalisk databas 182
lexikografisk metod 42
lyjdkälleverb 282

M

makrostruktur 267
manglende ekvivalens 82
Manifestationslogik 92
marknadsföring 261
maskinel oversættelse 29
materiale 248
megastruktur 265
mellantext 265
metaforik 227
metaforisk princip 285
middelnedertysk 123
mikrostruktur 204, 207, 268
mikrostrukturelle enheder 206
modersmål 170

morfem 183
morfografematiske ændringer 67
morfologisk variation 50
MT/HT for hhv. maskinel og traditionel
(human) oversættelse) 29
MT-system 30
Munksgaard 87
målgruppe 206

N

nationalordbøger 69
nivåer i framställningen 174
NLO 193
Nordisk lexikografisk ordbok 275
nordisk termbank 46
nordisme 80
norm 61, 126
normativ ordlista 41
normert terminologi 79
nukleært verb 281
nyskapad term 42
nøgleordskonkordans 35

O

oprecis term 42
ORACLE SQL 91
ord og sag 249
ordbogens funktion 206
ordbogsartikel 147, 148
Ordbogsmoduliet 30
ordboksredaktion 257
ordboksredaktör 257
ordklassetagging 67
ortografiske varianter 67
overgenerering 34, 66
oversættelse 78

P

parafraseangivelse 209
parsing 94, 130
partiell ekvivalens 83
passiv ordlista 41

perspektivmarkeringer 84
pladssymbol 210
polysemi 227, 279
polysemiindeks 207
pragmatiske ekvivalensrelasjoner 84
prepositionsuttryk 43
produktiva konstruktioner 47, 173
produktutveckling 261
projektgrupp 258
projektledare 258
projektplan 259
prototypisk betydelse 282
præpositioner 143, 147

R

realdefinitioner 252
redaktionsvejledning 87
reaktion 145, 149
relation 62
råmanuskript 109

S

saglig indgang til sproget 251
saglige citater 251
saglige definitioner 252
sagordbog 205
samform 85
sammansättning 167
sammensætninger 32, 67, 72
sekvenseringsenheder 207
semantisk valens 102
semantiska fält 279
semantiskt fält 104
SGML 95, 237
SGML-grammatik 132
SGML-standard 130
sidstebeleg 58
situationsbunden variation 51
slå- og stavfejl 67
specialordbøger 150
sprogbrug 61
sprogordbog 205

språkprov 168
stil 81
stilvalör 168
struktureret tekst 130
strukturindikatorer 207
superlativ 67
supplement 69
synonym 168
synonymi 79
synonymiangivelse 209
syntaktisk 171
syntaktisk valens 103
syntaktisk variation 50
systemarkitekturen 30
systematik i fackordbøger 263
systematisk introduktion 275
søgefelt 210

T

talesätt 49
tekstkorpus 61, 71
term 41
termbank 38
termharmonisering 264
terminolog 44
terminologi 38
terminologisk metod 42
terminologisk systematik 263, 273, 276
terminventering 264
termistor 266
termnummer 270
termpostformat 39, 45
termsystematik 263
termsystematisering 264
termval 44
termvalg 79
tertium comparationis 125
tilbagedatering 55
tillfälliga ordkombinationer 41
tosproget fagordbog 204
transem 78
transparens 80

træstruktur 131
tvärspråklig 227

U

undergenerering 66
universalordbog 150
universalordbøger 144
uppslagsord 167, 168
usus 66, 80
utanförtext 265
uttal 167
utvidgad 169
utvidgning 169
über 145, 146, 148

V

valens 144, 147, 149
valensskala 102
vending af ordbøger 237

W

WordPerfect 5.1 91
WYSIWYG 95

Æ

ækvivalent 205
ækvivalentposition 209

