



Dansk begrebsordbog: baggrund og metode

Lars Trap-Jensen

A Danish thesaurus project is currently underway at the Society for Danish Language and Literature in Copenhagen. Funded by The Carlsberg Foundation, the project runs 2010-2013 with a staff of two full-time editors. It does not start from scratch but draws on already existing lexical data: main input for the thesaurus will be the semantic descriptions of The Danish Dictionary (DDO, 116,000 senses) and the Danish WordNet (DanNet, 65,000 senses), the latter resource itself built on DDO data. Following an initial analysis of existing thesauri, a macro-structure was adopted close to the one used by Franz Dornseiff: *Der deutsche Wortschatz nach Sachgruppen* (2004). As a side benefit, this choice makes it easier for the editors to utilize the contents of the only existing thesaurus for Danish, Harry Andersen's *Dansk Begrebsordbog* (1945), which itself is no less structurally dependent on the first edition of Dornseiff (1934). One methodological aspect is given particular attention: what should be the main ordering principle within an entry? A final decision has yet to be taken as it is primarily a problem of presentation with no immediate bearing on the editing process, but eventually a conclusion must be reached as the thesaurus is going to appear as a print dictionary, too.

1. Baggrund

På Den Danske Ordbog gjorde vi for snart en del år siden den samme erfaring som alle andre der arbejder med elektroniske ordbøger, nemlig at ophævelsen af papirets alfabetiske ordningsprincip åbner for andre måder at systematisere materialet og skabe nye forbindelser mellem ord og betydninger som i forvejen er indeholdt i ordbogen. En følge af denne erfaring var udarbejdelsen af






et dansk wordnet i samarbejde med Center for Sprogteknologi, DanNet (jf. www.wordnet.dk, Pedersen m.fl. 2009). Undervejs i arbejdet med DDO havde vi gjort os nogle – på daværende tidspunkt ikke helt klare – forestillinger om at data kunne udnyttes på den måde, og vi havde derfor indført et felt i ordbogsstrukturen til angivelse af nærmeste overbegreb. Elementet var ikke synligt i den trykte ordbog, men kom os til gode ved etableringen af DanNet fordi det blev en genvej til at etablere hierarkiet af over- og underbegreber, som jo er en central del af et wordnet. Et andet ikke umiddelbart synligt element der har kunnet udnyttes, er et felt hvori redaktørerne skulle angive en betydnings tilknytning til et fagområde hvis det overhovedet var muligt – til forskel fra den mere snævre angivelse af fagbetegnelse. Dette har gjort det muligt at knytte semantiske domæner til mange betydninger i DanNet.

Et punkt hvor ordbogens oplysninger viste sig mindre hensigtsmæssige for DanNet, var i dens – i hvert fald i computersammenhæng – utilstrækkelige systematiske konsekvens. Det er som bekendt helt nødvendigt at være både eksplicit og konsistent i angivelse af de relationer der udvælges i beskrivelsen, når resursen efterfølgende skal kunne bruges af maskiner og i softwareapplikationer. Derfor er de eksisterende data blevet beriget i forhold til udgangspunktet så der ikke opstår tilfældige huller i kodningen.

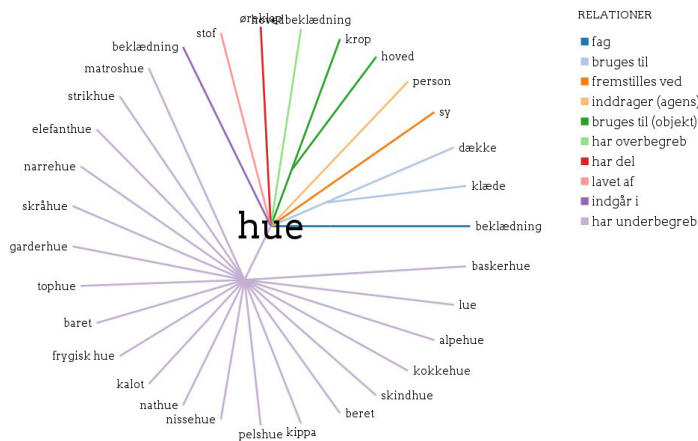
I slutfasen af DanNet-projektet har vi forsøgt at udnytte disse data også for menneskelige brugere. I Den Danske Ordbogs netudgave er det sket med oplysningstypen “Beslægtede ord”. Figur 1 viser hvilke beslægtede ord der er registreret for *hue* i Den Danske Ordbog. Som hovedregel vises der tre typer af beslægtede ord, hentet fra DanNet: overbegreber (mere generelle ord), underbegreber (mere specifikke ord) og søsterbegreber (ord der deler samme overbegreb, kohyponymer). Der er også mulighed for at se de beslægtede ord i en grafisk visning når man klikker på “vis som grafik (eksternt link)”. Resultatet kan ses i figur 2 (læs nærmere om denne funktion i Johannsen og Pedersen 2011).



BESLÆGTEDE ORD^{NETA} mere generelt: hovedtøj | hovedbeklædning || mere specifikt:
 elefanthue | lue | skråhue | kippa | baskerhue | kokkehue | skindhue | tophue |
 baret | alpehue | narrehue | beret | frygisk hue | kalot | nathue | nissehue |
 garderhue | strikhue | matroshue | pelshue || andre ord med "hovedtøj -
 hovedbeklædning" som overbegreb: hætte | hat | kyse | kappe | turban | kasket |
 bjørneskindshue | tiara | hjelm | sovsekande | bispehue | mitra | ørevarmer ...[skjul](#)
 vis som grafik (eksternt link) 

Figur 1: Beslægtede ord for hue i Den Danske Ordbog.

Den visuelle præsentation er dynamisk derved at man kan klikke på hvert af de beslægtede ord, hvorved figuren positionerer sig med det valgte ord som centrum og de dertil hørende relationer. Det fungerer ganske godt for nogle betydningsområder, især dem der opfører sig taksonomisk ordentligt, eksempelvis dette område hvor det drejer sig om konkrete genstande. I andre tilfælde fungerer det mindre godt. Det gør det fx ikke i tilfælde hvor hierarkiet har uoverskueligt mange medlemmer på hvert niveau. Den situation er ikke usædvanlig.



Figur 2: Visuel fremstilling af oplysningerne i figur 1.

I et tilfælde som *filminstruktør* nævnes der stort set kun søsterbegreber til filminstruktør, dvs. ord der alle betegner personer, og



TEMATISKE BIDRAG

dem er der for mange af til at det er overskueligt og brugbart for en menneskelig bruger. For at hjælpe brugerne en smule har vi derfor prøvet at skabe en ikke-tilfældig rækkefølge således at de formodet bedst egnede beslægtede ord vises først. Det er sket på grundlag af en score som beregnes efter hvor mange relationer der ud over overbegrebet binder to betydninger sammen, fx om de tilhører det samme fagområde, om de peger på den samme betydning uden selv at være forbundne osv.

BESLÆGTEDE ORD mere generelt: individ | person | hoved | menneske | mand | m/k'er || andre ord med "individ - person - hoved - menneske - mand - m/k'er" som overbegreb: operatør | stuntman | filmklipper | dekoratør | animator | filmoperatør | manuskriptforfatter | drejebogsforfatter | biografdirektør | klipper | biografefjer | biografgænger | instruktør | bedemand | kurator | auktionarius | arbejdsformidler | afdelingschef | impresario | administrator | forsikringsmægler | frimærkehandler | kapelmester | orkesterleder | dirigent | børnehavetilsættelse | udråber | forligsmand | journalist | skriverkarl | underviser | balletchef | belysningsmester | korrespondent | tasteoperatør | pædagog | logopæd | lærer | afdelingsleder | gæsteforelæser ...vis 160 flere

Figur 3: Beslægtede ord for filminstruktør i Den Danske Ordbog.

Det gør at personbetegnelser inden for domænet 'film' nævnes først (*operatør, stuntman, animator, filmoperatør, filmklipper, dekoratør, manuskriptforfatter, drejebogsforfatter, biografdirektør, klipper, biografgænger ... osv.*), hvad der vel hjælper lidt, men alligevel ikke er helt overbevisende.

Som det er fremgået, er over-underbegrebs-relationen den vigtigste når der skal findes beslægtede ord, og det er den væsentligste grund til at systemet fungerer bedst for konkrete substantiver. Hvis vi ser på en anden begrebstype, abstrakter, er det tydeligt at taksonomien er mindre veludbygget. Der er ikke så meget at sige taksonomisk om abstrakte handlinger: De har ét endnu mere abstrakt overbegreb 'handling' og derudover blot en mængde søsterbegreber som ikke behøver have alverden tilfælles i øvrigt. Et eksempel som *vold* resulterer således i:

ledelse, bevarelse, svinkeærinde, kunstgreb, gengæld, gestus, indledning, ridt, joke, tjeneste, fejltrin, fejltagelse, snuptag, slyngelstreg, skurkestreg, modtræk, gunstbevisning, fejlgreb,



misgreb, hævnakt, hævn, opfølgning, followup, dækmanøvre, beslutning, afgørelse, tyveri, behandling, gemenhed, akt, uanstændighed, indgreb, overraskelsesangreb, idræt, fedteri, operation, idioti, venlighed, judaskys, forbrydelse, ugerning, hug, modsvær, aktion, afværgemanøvre, rævestreg, galmandsværk, angrebsbølge, nederdrægtighed, camouflage, afledningsmanøvre, forvandlingsnummer, harakiri, sabotage, hyldest, ceremoniel, fingerknips, genistreg, ... (reduceret fra 205 forekomster)

De fleste vil nok bekræfte at man i dette tilfælde ville foretrække dels færre ord, dels ord fra samme semantiske felt, fx *magt, magt-anvendelse, gevalt, herredømme, overgreb, overlast, voldshandling, voldtægt, voldtægtsforbrydelse, voldførelse, brutalitet*. En indikator for at de anførte beslægtede ord kun er svagt beslægtede, ses af at den omtalte score ikke er særlig høj: 10 for de førstnævnte søsterbegreber, mens den er 110 for de første søsterbegreber nævnt under *hue*. Det samme problem gør sig gældende inden for ordklasserne verber og adjektiver: På den ene side får man for mange beslægtede ord, og i andre tilfælde mistænker man at relevante begreber ikke kommer med. Det skyldes formentlig dels at et wordnet og en begrebsordbog ikke har de samme primære anvendelsesmuligheder, dels at resurserne i DanNet ikke har rakt til den nødvendige manuelle finjustering af de semantiske hierarkier. En kritik man ikke så sjældent hører fremført mod wordnet, er at de er for fine i deres semantiske beskrivelser. I sprogteknologiske applikationer vil det ofte være en fordel hvis der om substantiver blot oplyses fx om det er en person og hvilken ontologisk type det tilhører. I DanNet har vi gjort en pointe ud af kun at operere med mellemniveauer hvis der findes ord i sproget for dette niveau. Pointen er at wordnettet skal afbilde det sproglige inventar i dansk og ikke være en begrebsontologi der står uden for sproget, sådan som man kender det fra fx faglige terminologier.



En anden – og beslægtet – forskel på wordnets og begrebsordbøger er at det i wordnetverdenen hører til sjældenhederne at en betydning indgår i flere begrebskategorier på samme tid. Det skal ikke forstås sådan at det er forbudt; det er det på ingen måde, og i DanNet forekommer det da også at et synset har flere overbegreber. Ofte er det fremkaldt af at der i en definition i DDO har været nævnt flere overbegreber. Fx defineres *datter* i DDO som en “pige eller kvinde som man er far eller mor til”, og derfor er *datter* blevet tilordnet både overbegrebet ‘pige’ og overbegrebet ‘kvinde’. Det kan også være at DanNet-redaktøren har besluttet at en betydning tilhører to overbegreber af systematiske grunde. Det er sket i tilfældet *blomkål*, som dels er placeret under ‘planter’, dels under ‘grøntsager’. Sproget indeholder mange af den slags eksempler uden at det i alle tilfælde er blevet udskilt som selvstændige betydninger i DDO. Men selvom det altså er tilladt at kode flere overbegreber for samme synset, er det ikke noget der er taget i anvendelse specielt ofte. I den nuværende version af DanNet er der lige under 500 tilfælde af synsets med flere overbegreber – ud af et samlet antal på 65.000 synsets.

Hvis man derimod slår op i registerdelen i en begrebsordbog, vil man hurtigt opdage at det er reglen snarere end undtagelsen at et opslag er placeret i flere kapitler. I den forbindelse skal man naturligvis huske at tage højde for at registeret ikke kan skelne mellem form og indhold, men kun har ordets betegnelse til rådighed og derfor ikke sondrer mellem homografer og forskellige ordbetydninger. Det er derfor begrebet ‘synset’ er vigtigt i wordnetverdenen, mens der ikke findes en pendant i almensproget. Men selv hvis man tager højde for den forskel, er det almindeligt i begrebsordbøger at samme ord i samme betydning (altså det der svarer til et synset med ét medlem) er placeret i flere forskellige kapitler i begrebsordbogen – som en afspejling af at ord og betydninger normalt ikke entydigt kan tilordnes ét og kun ét betydningsområde. Det er den væsentligste årsag til at begrebsordbøgers indekser





indeholder færre ord end der optræder i kapitlernes systematiske del (jf. opgørelserne i Nimb 2011).

Vi kan se på et enkelt eksempel: I DDO er der bare én betydning af *køb* som verbalsubstantiv ‘handlingen at købe’ (foruden en konkret betydning: ‘det man køber’, dvs. = vare). I DanNet er der angivet ét overbegreb (‘anskaffelse’) og ét underbegreb (‘import’), mens Harry Andersen 1945 anfører *køb* i tre forskellige indgange: 28 *Ombytning*, 808 *Handel* og 811 *Køb*. Det er den samme betydning der optræder alle tre steder, men der er nuanceforskelle alt efter om man har fokus på købet som et bytteforhold, på handel i almindelighed eller på den specifikke købsituation. Det er der ikke noget usædvanligt i; det er snarest den almindelige situation når menneskelige brugere griber til en begrebsordbog: Vi kan opfatte den samme betydning som hørende til forskellige områder, alt efter hvilken synsvinkel vi anlægger. Og det er også derfor jeg tidligere har karakteriseret det som problematisk når en betydning kun tilordnes én kategori i en tesaurus (jf. fx Trap-Jensen 2010:300ff.). Hvis man fx slår *guinea pig* op i Macmillans netordbog, optræder marsvinet kun ét sted, nemlig i en kategori der benævnes: ‘Animals found on more than one continent’, og det er langt fra sikkert at det er den kategori brugeren er interesseret i. Lige så sandsynligt er det at brugeren interesserer sig for ‘kæledyr’, ‘gnavere’ eller måske ‘dyr med pels’, ‘små pattedyr’ eller ligefrem ‘dyr der kan spises’. Pointen er at det er meningsfuldt at placere marsvinet i flere forskellige kategorier – der er ikke altid kun én relevant kategori for en bestemt betydning.

En vigtig opgave består derfor i at forbedre funktionen Beslægtede ord ved at oprette flere meningsfulde kategorier med et overskueligt antal relevante medlemmer i hver. Det var ikke mindst dette der motiverede os til at gå i gang med en egentlig begrebsordbog.





2. Metode

2.1. Makrostruktur

I sammenligning med andre begrebsordbøger er vores projekt et discountprojekt: 3 år med en bemanning på 2 redaktører er ikke meget i forhold til de livstidsprojekter som de fleste forgængere har været. Til gengæld begynder vi ikke fra bunden, men har et omfattende og systemiseret materiale som udgangspunkt. I en vis forstand drejer det sig om finpudsning af det eksisterende materiale og en præsentation af det i en form der er hensigtsmæssig for ordbogsbrugere.

Ét sted hvor vi har valgt en nem løsning, er med hensyn til makrostrukturen. I stedet for at gentænke en konceptuel opdeling af verden, har vi valgt at overtage en eksisterende i stedet. Mest oplagt ville det nok være at bruge Rogets opdeling i ca. 1000 (2002-udgaven: 990) grupper. Ikke alene er den med over 40 millioner solgte eksemplarer blandt de 3 mest sælgende ordbøger i historien (ifølge Wikipedia kun overgået af Websters American Spelling Book (100 mio.) og Merriam-Websters Collegiate Dictionary, 55 mio.), men vigtigere endnu er den pga. sin store kommercielle succes kommet i et utal af nye udgaver og revisioner i løbet af de over 150 år der er gået siden den udkom første gang. Af samme grund har Roget været forbillede og udgangspunkt for de fleste begrebsordbøger der siden er udkommet. I dansk leksikografi finder vi klassikeren Harry Andersens *Dansk Begrebsordbog* fra 1945, og den lægger sig igen ganske tæt op ad den tyske klassiker, Franz Dornseiffs *Der deutsche Wortschatz nach Sachgruppen*. På tidspunktet for udgivelsen af Andersen 1945 var Dornseiff en forholdsvis ny udgivelse, idet den udkom første gang i 1934. Dornseiff er også inspireret af Roget, men den kopierer ikke strukturen fra Rogets. I stedet opererer den med 22 større kapitler der igen er inddelt i 906 tematiske grupper. Det mærkes at Dornseiff er udkommet over 80 år senere





end Roget: Den er ikke på samme måde forankret i romantikkens opfattelse af sprogets organiske sammenhæng og traditionen fra Leibniz og bagud til Aristoteles. Dornseiff er for tysk nærmest blevet hvad Roget er for engelsk, og den er på tilsvarende vis kommet i mange senere revisioner, senest i en 8. udgave i 2004 med cd-rom.

I Det Danske Sprog- og Litteraturselskab er vi nået til samme konklusion som Harry Andersen og har besluttet at lægge os op ad makrostrukturen hos Dornseiff. Dens opdeling af verden virker mere klar og mere intuitivt gennemskuelig end Rogets ældre inddeling. Det har den yderligere fordel at det er nemt for os at supplere med relevant materiale fra Harry Andersens begrebsordbog.

I stedet for at genopfinde den dybe tallerken kopierer vi derfor som udgangspunkt makrostrukturen fra Dornseiff, hvorefter alt indhold importeres fra DanNet og DDO, jf. Nimb (2011). Derved får vi også hele mediostrukturen forærende. En vigtig del af en begrebsordbog er jo – som vi så før – at der dels oplyses i indekset om alle de relevante grupper et ord står i, dels at der henvises fra én tematisk gruppe til andre betydningsbeslægtede grupper fordi samme ord eller betydning kan optræde i forskellige grupper alt efter hvilken synsvinkel det betragtes under.

Vi indrømmer gerne at det er en discountløsning. De uhen-sigtsmæssigheder der måtte være i Dornseiff og Harry Andersen, importeres automatisk til vores begrebsordbog. Den risiko følger med, men det samme gælder naturligvis også den sikkerhed der ligger i de revisioner og tilføjelser der løbende er foretaget frem til 8. udgave, som er vores udgangspunkt. Og man er nødt til at sammenholde det med de resurser der skulle have været investeret i at koncipere en alternativ struktur. Vi tror ikke det ville føre til et væsentlig bedre resultat end det foreliggende i Dornseiffs 8. udgave. Og det ville være ærgerligt at bruge mange resurser på at nå frem til et resultat der måske ikke ville afvige meget fra det foreliggende. Så det har egentlig været en nem beslutning at tage.

Løsningen fører selvfølgelig det væsentlige spørgsmål med sig:





I hvilken udstrækning lader den verdensopdeling som Dornseiff bruger, sig overføre fra et sprog til et andet? Kan man være sikker på at de grupper der byder sig til på tysk, også gør det på dansk? Nej, det kan man ikke; begreber er naturligvis ikke kongruente på tværs af sprog. Uden at gå helt tilbage til Sapir-Whorf og sammenhængen mellem sprog og tanke turde det være klart at der er både sproglige og kulturelle forskelle som man skal være opmærksomme på. Vi kan se på et par eksempler.

Dornseiffs gruppe 17.24 hedder *Spiele*. Det er formentlig uproblematisk på tysk at gruppere *Handpuppe*, *Eisenbahn* og *Schaukelrferd* i samme gruppe som *Monopoly*, *Schach* og *Bridge* fordi man på tysk har ét fælles begreb og ikke skelner – som vi gør på dansk – mellem ‘spil’ og ‘leg’.

Det kan også gå den anden vej; at der er to tyske begreber, hvor vi på dansk kun vil have ét. Og endelig er der alle de kulturbetingede forskelle mellem en dansk og en tysk virkelighed som slår igennem i begrebsopdelingen: Dornseiff har fx et selvstændigt afsnit om industri og har tillige udskilt to særlige områder, bilindustri og medicinalindustri, i selvstændige afsnit. Her ville det i en dansk sammenhæng være mindst lige så oplagt at udskille fx fødevarerindustri eller turisme. Og Dornseiff har et stort afsnit 18.10 *Verwaltungsbehörde* ved siden af et andet afsnit 18.11 om *Vertretungsausschuss*, *Ratsversammlung*. De mange begreber der afspejler tysk samfundsstruktur på forbunds- og delstatsniveau, har ikke nødvendigvis sidestykke i en dansk sammenhæng.

Dansk og tysk er altså ikke kongruente i deres konceptualisering af verden. Alligevel er det vores vurdering at det samlet set drejer sig om en beskeden del af begrebsgrupperne. Det bliver derfor nødvendigt at foretage visse redaktionelle tilpasninger i makrostrukturen efterhånden som behovet opstår: Nogle grupper skal slås sammen, og andre opdeles i flere. Redaktørerne skal i den forbindelse være opmærksomme på de ændringer i henvisnings-systemet det medfører.





2.2. Mikrostruktur

En enkelt beslutning har redaktionsinternt voldt særlige problemer. Det drejer sig om præsentationen af data på det øverste niveau i mikrostrukturen.

Der lader til at være enighed mellem de fleste eksisterende begrebsordbøger om følgende prioritering af kriterier for ordning af mikrostrukturen:

1. ordklassebestemt
2. semantisk: afsnit – gruppe – undergruppe
3. alfabetisk

Hos Roget er en artikel først opdelt i ordklasser: Substantiv, adjektiv, verbum og interjektion. Hver ordklasse er underinddelt i mindre grupper bestemt af betydningsnuance og placeret i selvstændige afsnit indledt af et nøgleord i kursiv. Internt i gruppen er inventaret ordnet i mindre grupper adskilt af semikolon og med begreberne anført i logisk rækkefølge. I slutningen af sådan en semikolongruppe kan der være henvisninger til andre nøgleord, anført i kursiv og med det pågældende nummererede afsnit.

Hos Dornseiff anvendes en mikrostruktur der minder om Rogets, men alligevel adskiller sig. Henvisninger optræder ikke inde i de enkelte grupper, men er placeret indledningsvis i en selvstændig rubrik. Ordklasseopdelingen er ikke eksplicit, men er dog også her det primære kriterium for inddelingen i mikrostrukturen, med den forskel at interjektioner kommer først, derefter substantiver, adjektiver/adverbier og til sidst verber. Ved ikke at have eksplicite ordklasseangivelser undgår Dornseiff klassifikationsproblemer inden for systemet, fx når det drejer sig om at placere præpositionsforbindelser, talemåder, rutineformler osv. Til gengæld er det sværere for brugerne at orientere sig i det enkelte afsnit.

Inden for den enkelte gruppe i Dornseiff er der plads til be-





tydningsnuancer. De er adskilt af en større firkant, mens de enkelte medlemmer i undergrupperne adskilles af en midtstillet prik. Grupper adskilt af store firkanter svarer til Rogets paragraffer. Også hos Dornseiff har det første ord i hver af de større grupper karakter af nøgleord, blot er det ikke kursiveret som hos Roget. Men dets status som nøgleord viser sig ved at det står uden for den alfabetiske rækkefølge som ellers er det ordnende princip internt i hver undergruppe.

Harry Andersens artikelstruktur minder i sagens natur mest om Dornseiffs, dog er brugen af henvisninger i en artikel mindre hyppig, og hvis der bruges henvisninger, kan de stå både på artikelniveau og i de enkelte grupper. Ordklasseopdelingen er gjort eksplicit, men der forekommer kun tre ordklasser, benævnt S., A. og V. Mens S. for substantiver og V. for verber er forholdsvis uproblematisk, fungerer A. som en restkategori. Den dækker over adjektiver, adverbier, pronominer, præpositioner, konjunktioner, interjektioner og talord – foruden fraselignende enheder: “til at brække sig over”, “Gud være lovet” osv. De enkelte undergrupper svarer i øvrigt til Dornseiffs, men adskilles af en tankestreg, og Harry Andersen bruger ikke nøgleord, men alene alfabetisk ordning af begreberne i hver undergruppe.

Selvom der altså er mindre forskelle i fremstillingen i de tre værker, er der enighed om det centrale spørgsmål: Ordklasserne er det vigtigste ordningsprincip. Man må derfor spørge: Er det nu så selvfølgelig? Derved kommer man jo til at adskille ord der står hinanden semantisk nær, blot fordi de tilhører forskellige ordklasser.

Svaret på det spørgsmål har ifølge Wiegand (2004) med konsultationssituationen at gøre. Han skriver herom:

Bei der oben formulierten Charakterisierung des genuinen Zwecks von großen einsprachigen Sachgruppenwörterbüchern wurden bei den Möglichkeiten, die Wörterbücher



dieses Typs dem Benutzer eröffnen, drei verschiedene Arten [unterschieden:

- (i) Konsultationssituation wegen Ausdrucksfindungsschwierigkeiten
- (ii) Konsultationssituation wegen äquivalentbezogener Angemessenheitszweifel
- (iii) Konsultationssituation wegen Wissensbedarf über einen bestimmten Aspekt des deutschen Wortschatzes]

Die erste Möglichkeit ... wird wahrscheinlich am meisten genutzt. Das ist auch der Grund dafür, warum die Wörterbuchartikel im Formteil nach Ausdrucksklassen und damit auch nach Wortarten gegliedert sind. Die kundigen Benutzer, die Ausdrucksfindungsschwierigkeiten haben, werden durch diese artikelinterne Datenanordnung präferiert.

Wiegand, 2004: 59 (i-iii sammenfattet fra 57-58)

Jeg er ikke bekendt med undersøgelser der empirisk har registreret brugen af begrebsordbøger, men Wiegands opfattelse harmonerer godt med min egen fornemmelse: Begrebsordbøger er først og fremmest produktionsordbøger der bruges til fremfinding af paradigmatiske alternative ord og udtryk i en bestemt sproglig kontekst. Kun sekundært bruges den til at skaffe overblik over sprogets inventar inden for et bestemt semantisk felt, og det samme gælder formentlig oversættelsessammenhængen, der til nød kan opfattes som et undertilfælde af (i). Man kan naturligvis indvende at alternative formuleringsvarianter ikke altid har form af paradigmatiske ledalternativer, og at det søgte ord eller udtryk derfor heller ikke nødvendigvis tilhører samme ordklasse. Dette synspunkt har utvivlsomt noget for sig, men spørgsmålet er hvor stor vægt det kan tillægges.

Man kan også have den opfattelse at (iii) skal prioriteres, altså



at det er vigtigere at strukturen prioriterer semantikken således at de enkelte undergrupper opregner alle ord inden for samme semantiske nuance uanset hvilken ordklasse det drejer sig om. Det har yderligere den fordel at det redigeringsteknisk er enklere eftersom det eksisterende materiale i forvejen er klassificeret efter semantiske kriterier, jf. Nimb (2011). Som det også nævnes dér, er det dog primært et præsentationsproblem som principielt er uafhængigt af redigeringsprocessen. Denne foregår først og fremmest efter semantiske principper, men alligevel kan indholdet forholdsvis enkelt præsenteres ordklasseopdelt i den færdige publikation. Det siger sig selv at valget først og fremmest er relevant når det drejer sig om en papirudgivelse. Begrebsordbogen skal dels udkomme som en traditionel ordbog på papir, dels indgå i netudgaven af Den Danske Ordbog til forbedring af den onomasiologiske komponent. Hvordan redigeringen af de enkelte kapitler finder sted, beskrives i Nimb (2011).

Litteratur

Ordbøger

- Andersen, Harry 1945: *Dansk Begrebsordbog*. København: Munksgaard.
- DDO = Ebba Hjorth, Kjeld Kristensen m.fl. 2003-2005: *Den Danske Ordbog*. København: Gyldendal og Det Danske Sprog- og Litteraturselskab. Onlineversion: ordnet.dk/ddo (februar 2011).
- Dornseiff, Franz 2004: *Der deutsche Wortschatz nach Sachgruppen*, 8. Auflage. Berlin/New York: Walter de Gruyter.
- Macmillan = Macmillan Online Dictionary. www.macmillandictionary.com
- Roget, Peter Mark: *Roget's Thesaurus*, 150th anniversary edition edited by George Davidson 2002. London: Penguin.



Anden litteratur

- Johannsen, Anders & Bolette S. Pedersen 2011: "Andre ord" – a wordnet browser for the Danish wordnet, DanNet. I: Proceedings from 18th Nordic Conference of Computational Linguistics, NODALIDA 2011, Riga, Latvia. Northern Association for Language Technology, University of Tartu, Vol. 11, 295-298.
- Nimb, Sanni 2011: Semantiske relationer i en ny dansk begrebsordbog: genbrug på tværs af ordbøger. I: *LexicoNordica 18* (i dette bind).
- Pedersen, Bolette S., S. Nimb, J. Asmussen, N. Sørensen, L. Trap-Jensen, H. Lorentzen 2009: The challenge of compiling a WordNet for Danish by reusing a monolingual dictionary. I: *Language Resources and Evaluation, Computational Linguistics Series*, Volume 43, Number 3: 269-299.
- Trap-Jensen, Lars 2010: Access to Multiple Lexical Resources at a Stroke: Integrating Dictionary, Corpus and Wordnet Data. I: Sylviane Granger, Magali Paquot (eds.): *eLexicography in the 21st Century: New Challenges, New Applications. Proceedings of eLex 2009, Louvain-la-Neuve, 22-24 October 2009*, Presses universitaires de Louvain • Cahiers du CENTAL 2010, 295-302.
- Wiegand, Herbert Ernst 2004: Lexikographisch-historische Einführung. I: Dornseiff 2004, 9-91.

Internethenvisninger

andreord.dk = www.andreord.dk ved Anders Johannsen. Viser data fra DanNet.

DanNet = Pedersen, Bolette S. m.fl. 2011: *DanNet*. www.wordnet.dk

Lars Trap-Jensen
ledende redaktør, cand.mag., MPhil.
Det Danske Sprog- og Litteraturselskab
Christians Brygge 1
DK-1219 København K
ltj@dsl.dk