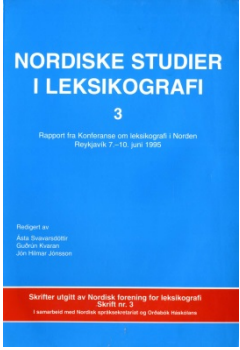


NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Tekstbaseret lemmaselektion	
Forfatter:	Grete Duvå og Anna-Lise Laursen	
Kilde:	Nordiske Studier i Leksikografi 3, 1995, s. 93-105 Rapport fra Konferanse om leksikografi i Norden, Reykjavík 7.-10. juni 1995	
URL:	http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive	

© Nordisk forening for leksikografi

Betingelser for brug af denne artikel

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

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

Søgbarhed

Artiklerne i de ældre Nordiske studier i leksikografi (1-5) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

Tekstbaseret lemmaselektion

Based on a research project at The Aarhus School of Business, which includes a text corpus of about 1 800 000 words established in preparation for lexicographical processing, problems regarding the use of a text corpus for lemma selection are discussed. Among other things, the priority of a text corpus compared to that of other sources in this phase is discussed, and an efficient search method used as an alternative to traditional selection methods is suggested.

1 Oprettelse af tekstkorpus

1.1 Baggrund

I 1991 og 1992 blev der på Handelshøjskolen i Århus gennemført undersøgelser, der fokuserede på brugen af ordbøger. Formålet med disse undersøgelser var bl.a. at kortlægge de problemer, som opstår under oversættelse af fagtekster samt at få viden om, hvordan oversætterne løser problemerne, og om hvilke hjælpemidler de benytter sig af.

Resultaterne fra analysen af brugerundersøgelserne, som bl.a. er beskrevet i Duvå/Laursen/Maidahl (1991), blev bl.a. brugt som udgangspunkt for udarbejdelse af forslag til en forbedring af den eksisterende tosproglige fagleksikografi. Forslagene tog hensyn til de behov, som undersøgelserne viste, at oversætterne havde.

For at imødekomme disse behov skulle en ideel fagordbog bl.a. indeholde definitioner, kollokationer og eksempler. Derudover skulle den kunne formidle fagindsigt til brugerne, da undersøgelserne viste, at oversætternes største problem var manglende faglig viden, og vi hævder, at uden denne viden og indsigt i det pågældende fagområde, kan der ikke produceres oversættelser af tilfredsstillende kvalitet. I de nævnte undersøgelser kunne oversætterne ikke hente hjælp til erhvervelse af denne viden i de eksisterende ordbøger, idet ordbøgerne ikke indeholdt informationer om sagen eller de såkaldte encyklopædiske oplysninger.

Vi besluttede på denne baggrund at undersøge, om et tekstkorpus kunne være et nyttigt redskab for den tosprogede fagleksikografi, dvs. om et sådant korpus via bearbejdning kunne danne basis for generering af en specialordbog inden for regnskabsvæsen, som kunne formidle brugerne den fagindsigt, som de ifølge vore brugerundersøgelser manglede. Vi vil i det følgende diskutere korpus egnethed i forhold til andre typer materiale i forbindelse med især lemmaselektion.

Vi mener, at det er af stor betydning for brugeren, at fagordbøger på det mikrostrukturelle plan indeholder ovennævnte oplysninger, men vi mener derudover, at opbygningen af ordbogens makrostruktur spiller en væsentlig rolle i denne forbindelse. Således argumenterer vi i det følgende for, at en fagordbog skal være tematisk eller onomasiologisk opbygget, da vi mener, at en systematisk makrostruktur umiddelbart vil kunne bidrage til at formidle en del af den faglige indsigt, som en fagoversætter har behov for (cf. Christiansen/Duvå/Laursen 1994). Vi går således et skridt videre end Pedersen (1995), der argumenterer for, at fagets

systematik ikke skal afspejles i makrostrukturen, men derimod præsenteres i ordbogens omtekst i form af faglig indledning eller lignende.

Det traditionelle ordbogskoncept med en semasiologisk opbygning bidrager ikke umiddelbart til at formidle faglig viden, idet en alfabetisk ordliste ikke giver brugeren information om de begrebsmæssige relationer mellem termene. Manglen ved en sådan alfabetisk liste skulle være evident. Her følger lemmata i fuldstændig vilkårlig rækkefølge, dvs. de forskellige faglige deluniverser blandes totalt sammen, og i bedste fald forsøger leksikografen at afhjælpe denne mangel ved en eller anden form for henvisningssystem til faglige informationer.

Eksempel på alfabetisk liste (Fryd 1990):

[...]
 entreprenørvirksomhed
 entreprise
equity metode = indre værdis metode
 erfare
 erfaringens vej, ad
 erholde

Et eksempel skal illustrere fordelene i denne forbindelse ved en systematisk opbygning frem for en alfabetisk (eksemplet er konstrueret):

- 1 konsolideringsformer
- 1.1 linie for linie konsolidering
- 1.1.1 (fuldstændig) konsolidering
- 1.1.2 pro rata konsolidering
- 1.2 konsolidering på én linie
- 1.2.1 **equity-metode**
- 1.2.1.1 eliminerings af interne fortjenester og tab

Fordelen ved den systematiske opbygning skulle være klar: en bruger, der skal oversætte fx *equity-metoden*, får umiddelbart begrebet præsenteret i sin begrebsmæssige sammenhæng, dvs. vedkommende får alene ved at se på lemmaudsnittet indsigt i, hvilket begrebsmæssige univers equitymetoden hører hjemme i. Er vedkommende regnskabsmæssig ukyndig får han/hun her straks en viden om, at det drejer sig om en konsolideringsmetode og kan ved at bevæge sig op eller ned i den systematiske lemmaliste supplere sin faglige viden alt efter behov. Den systematiske ordning forudsætter naturligvis en begrebsmæssigt udtømmende lemmabestand.

Vi vil i det følgende se på, hvilke typer materiale der kan komme i betragtning til generering af en sådan **systematisk lemmaliste** og senere se på, hvilke kriterier og metoder der kan anvendes i den forbindelse.

1.2 Ordbøgers materiale

Den empiriske basis kan bestå af følgende elementer cf. Bergenholtz/Tarp (1995), herefter betegnet Manualen, hvis opdeling vi — med visse ændringer — i store træk følger.

1. Egen fagsproglige og faglige kompetence eller viden om det pågældende område (i manualen kaldet introspektion — et begreb, som vi mener snarere refererer til metoden, cf. nedenfor)
2. Belægsamlinger
3. Foreliggende beskrivelser, der kan deles i a) ordbøger b) håndbøger, leksika, encyklopædier med flere c) lærebøger og monografier og d) videnskabelige artikler
4. Tekstkorpora

1.2.1 Egen fagsproglige og faglige kompetence

Om det er muligt at selekttere lemmata til en fagsproglig ordbog uden anden basis end leksikografens eget aktive og passive ordforråd inden for et givet fagsprogligt område, må i princippet afhænge af, i hvor høj grad den faglige kompetence er til stede. Hvis det handler om en ekspert på fagområdet, må man gå ud fra, at en så godt som udtømmende bestand af det til faget hørende register er til stede hos den pågældende, men som det anføres i manualen (p. 92), kan modersmålskompetencen være inkomplet. Dette er bl.a. tilfældet, hvis fx undervisning og publikation af artikler foregår på et andet sprog end modersmålet. En dansk ekspert i afsætningsøkonomi kan fx i sit aktive ordforråd rumme termen *skimming price*, uden at kunne gøre rede for, hvad begrebet hedder på dansk.

1.2.2 Belægsamlinger

Belægsamlinger er den klassiske basis for fagsprogsordbøger. Det drejer sig om eksemplarsamlinger, som fx oversættere — typisk translatører — via deres daglige arbejde har nedfældet på kartotekskort eller i en database. De udgør en slags erfaringsbank, og da de afspejler helt konkrete erfaringer fra de daglige arbejdsopgaver, er de særdeles relevante. På den anden side vil de som følge af en fagoversætters mangeartede oversættelsesopgaver som regel være tilfældigt udvalgte og vil derfor sjældent rumme et fuldstændigt register af termer inden for et konkret område. Og hvis en leksikograf, der samtidig er oversætter, oven i købet vælger ukritisk at publicere en komplet belægsamling af ovennævnte blandede art, bliver resultatet tilfældigt, metodisk kritisabelt, og undertiden ganske morsomt.

Vi har hentet nogle eksempler på en sådan uheldig lemmaselektion fra en dansk-spansk ordbog (Galindo 1989) der angiveligt skulle dække området *erhverv og økonomi*:

alfonseri	postbedrageri
bordel	svangerskabsforebyggelse
diamantbryllup	rufferi
posebæreevne	vielse
positionslys	

Sådanne belægsamlinger vil ofte afspejle leksikografens eget univers snarere end det univers, ordbogen angiver at dække.

1.2.3 Foreliggende beskrivelser

Under denne overskrift inkluderer manualen et vidt spektrum af fagtekster, herunder ordbøger, håndbøger, leksika samt lærebøger og relevante artikler.

At anvende ordbøger som eneste empiriske basis ville være kontradiktorisk, idet motivationen for at gå i gang med en ordbog ofte er manglende eller mangelfulde ordbøger på det pågældende område. Evt. kan eksisterende ordbøger fungere som inspirationskilde eller som *reminders*, det vil fx sige som kontrolmateriale for en lemmaselektion.

Anvendelse af indeks eller stikordsregistre til håndbøger og lærebøger som basis for en lemmaselektion og strukturering heraf kan virke relevant. Vi har imidlertid foretaget adskillige tjek inden for vores fagområde, der viser, at sådanne indeks er ufuldstændige og i alt fald kun dækker fagområdets øverste begrebsmæssige niveau. Hvis sådanne indeks anvendes, er det derfor nødvendigt at gå videre fra opslagsordet til selve teksten for at supplere bredden og dybden af lemmata. Den egentlige basis for selektionen vil derfor være **teksten** som sådan med indekset som det strukturerende udgangspunkt.

1.2.4 Tekstkorpus

Ved tekstkorpus forstår vi en samling tekster, der er produceret i en kommunikativ sammenhæng (cf. Bergenholtz/Pedersen 1994), og som desuden er specielt sammensat. Som sådan adskiller det sig fra belægsamlinger samt fra ordlister og ordbøger. Et sådant korpus bestående af en velovervejede samling af fagtekster mener vi, er det bedste udgangspunkt for arbejdet med fagsprogsordbøger.

Hvis man vælger tekster af forskellige forfattere og med forskellige funktioner (lovttekster, lærebøger, håndbøger mv.), der begrebsmæssigt dækker det pågældende fagområde, vil man være nogenlunde sikker på, at de fleste centrale termer i fagsystematikken bliver repræsenteret. Det skal tilføjes, at det ikke, som det er praksis for korpora, der er beregnet på syntaktiske analyser, er hensigtsmæssigt at tage stikprøver ud fra en større samling tekster. Eftersom det handler om et velafgrænset fagområde, der begrebsmæssigt skal dækkes helt, vil det rigtige være at vælge en række hele tekster, der repræsenterer termsystematikken i det pågældende fagområde.

I det tidligere nævnte ordbogsprojekt vedrørende regnskabsterminologi vil der således ikke være nogen ide i at tage et afsnit om balancen fra én håndbog og så vælge fx afsnit om noter, værdiansættelse og afskrivninger fra tre andre håndbøger. Det drejer sig derimod om at vælge et par anerkendte håndbøger, der formodes at have terminologisk gennemslagskraft, dvs. som fx bliver anvendt i uddannelsen af regnskabseksperter.

Ved sammensætning af sådanne korpora, dvs. ved bevidst at vælge nogle enkelte centrale og anerkendte værker inden for tekster med forskellig funktion, kan man ikke tale om, at korpus er repræsentativt i traditionel forstand, idet dette forudsætter stikprøveudtagning fra en kendt grundmængde (Manualen s. 95). Valg af typiske tekster, der „efter bedste skøn“ dækker det pågældende fagområde kaldes i Manualen for *eksemplarisk*, og vi kan tilslutte os denne skelnen.

Med hensyn til omfang af et sådant tekstkorpus nævner manualen, at hvor det til almenproglige ordbøger vil være nødvendigt med et korpus på 50 millioner tekstord, er det vanskeligere at afgrænse dette omfang for fagsprogskorpora. Der nævnes et konkret eksempel i Manualen på, at et korpus på 500 000 ord har været dækkende til en fagordbog.

1.3 Beskrivelse af *Regnskabskorpus*

1.3.1 Emneområde

Emneområdet for vores tekstkorpus er som nævnt regnskabsvæsen, i første omgang med hovedvægt på eksternt regnskabsvæsen.

Da der er tale om en kontrastiv bearbejdning (da-fr, da-sp) af tekstmaterialet, er der etableret tre delkorpora.

1.3.2 Sammensætning af korpus

Korpus er sammensat af følgende teksttyper, som er de centrale for området:

- love
- kommenterede lovtekster
- regnskabsstandarder
- regnskabsvejledninger
- autentiske regnskaber
- lærebøger
- artikler

Udarbejdelse af bibliografi og sammensætning af korpus er sket i samarbejde med fageksperter.

1.3.3 Omfang

Det danske delkorpus er på ca. 800 000 løbende ord, hvilket stort set svarer til lignende korpora. Om korpus i denne størrelsesorden er tilstrækkeligt til at afspejle områdets lemmata må bero på et skøn foretaget af leksikografen og fageksperter i fællesskab under hensyntagen til det givne formål.

I øvrigt mener vi ikke, at det er muligt at give præcise kriterier for, hvor stort et omfang et korpus bør have, idet der er mange faktorer, der spiller ind såsom formål, omfang af det faglige univers, man vil undersøge, ressourcer, lagerkapacitet etc. Men efter vort skøn er et korpus på omkring 1 mio. løbende ord et rimeligt udgangspunkt for et eksemplarisk fagkorpus.

Der er af komparative hensyn tilstræbt symmetri mellem delkorporaerne L1 og L2 på det tematiske og teksttypologiske niveau.

2 Fra korpus til lemmasektion

2.1 Kriterier for lemmasektion

Der kan opstilles forskellige kriterier for opstilling af en lemmaliste til en fagordbog. Vi mener, at følgende to hovedkriterier er relevante i forbindelse med generering af en systematisk specialordbog:

1. alene fagord, men derimod *alle* fagord uanset frekvens, der indgår i fagsystematikken, medtages, eller
2. fagord som beskrevet i 1) samt andre ord udvalgt efter andre kriterier, herunder verber, adjektiver, adverbier og orddele (affikser, ordstammer mv.). Andre kriterier kan fx være frekvens eller *pædagogiske* erfaringer for, hvad der almindeligvis volder problemer.

Vi mener ikke, at kun fagord/termer skal medtages, men at det af hensyn til brugers behov er væsentligt, at også andre ord, der er specielt relevante for det pågældende fagområde, medtages som lemmata. Men vi understreger, at de skal være specielt relevante. Vi mener således ikke, at almindelige økonomiske fællesord (som fx *pris, beregning, beregne, kontanter, købe, sælge, låne, lån*) skal medtages som lemmata, men kun de ord, der har en specifik betydning eller en specifik funktion inden for området regnskab, fx verber som *udgiftsføre, afskrive, konsolidere, værdiansætte, ansætte, optage til* [. . .].

I vores tematisk/onomasiologiske koncept vil vi i første fase selekttere termer som lemmata, og det betyder, at flerordstermer selekteres og lemmatiseres på lige fod med enkeltordstermer. Andre relevante ord vil vi senere lade indgå i ordbogen enten som fx kollokationer eller i særlige systematiske registre (fx verber systematiseret efter proces: *aktivere, værdiansætte, værdiregulere (nedskrive, opskrive, afskrive* [. . .]).

2.2 Metoder til lemmasektion

Ved selektion af lemmata til en fagsproglig ordbog, forekommer der *a priori* tre metoder:

1. Introspektion (her forstået som brug af den kompetence, leksikografen selv er i besiddelse af)
2. Multispektion (her forstået som et samarbejde mellem en eller flere fagsprogskyndige og en eller flere fagkyndige)
3. Automatisk sortering

Vi forudsætter, at den valgte empiriske basis er et til formålet sammensat tekstkorpus, idet vi som nævnt mener, at et sådant er det bedste grundlag for generering af en fagordbog. Den leksikografiske bearbejdning af tekstkorpus frembyder to åbenbare muligheder.

Man kan excerperere ved en manuel gennemgang af materialet, idet man ved en form for introspektion el. evt. ved multispektion afgør, hvilke lemmata der skal medtages. Denne proces er umådeligt tidskrævende.

Den anden mulighed er, at man indscanner teksterne og udskriver en alfabetisk ordliste, som leksikografen kan sortere ud fra. Denne fremgangsmåde har den ulempe, at listen kun indeholder enkeltord, således at man ved potentielle flerordstermer skal ind i teksten igen for at kontrollere ordets omgivelser. Dette kræver gode associerende evner hos leksikografen, hvilket følgende eksempel fra vort regnskabskorpus kan illustrere:

afskrive	afskrivninger
afskrivelige	afskrivningsberettigede
afskriveligt	afskrivningsfond
afskriver	afskrivningsforløb
afskrives	[...]
afskrivning	afskrivningsmetode
[...]	[...]
	afskrivningsperiode
	[...]
	afskrivningsplan

Blandt de anførte ord i listen er der flere, som vi på forhånd ved indgår i flerordstermer. *Afskrivning* fx indgår i *lineær afskrivning* og i *skattemæssig afskrivning*, *afskrivelig* indgår fx i *afskrivelig omkostning*. Men en sådan introspektion vil næppe nogensinde afsløre samtlige flerordstermini omkring en hovedterm og ej heller afdække alle de listede ord, der er kandidater til en flerordsterm. I alle tilfælde vil behovet for multispektion til bedømmelse af termsystematik og synonymi være til stede.

En fuldstændig automatisk sortering kan finde sted, hvis man sammenholder den alfabetiske lemmaliste med et korpus fra det almensproglige område/et andet fagsprogligt område. Vi foretog et sådant eksperiment i starten af vores projekt, hvor vi sammenlignede de danske regnskabsstandarder med et eksisterende korpus af almensproglige tekster. Man kan på denne måde få en koncentreret liste af fagtermer, men maskinen sorterer på den anden side mange termer fra, der indgår i både det almensproglige og det specifikke fagsproglige univers. Der kan undertiden være tale om homonymer, og i andre tilfælde fagord, der er blevet optaget i almensproget. Eksempler på sådanne termer, der blev sorteret fra som alment ordforråd var *aktiv*, *post* og *skat*.

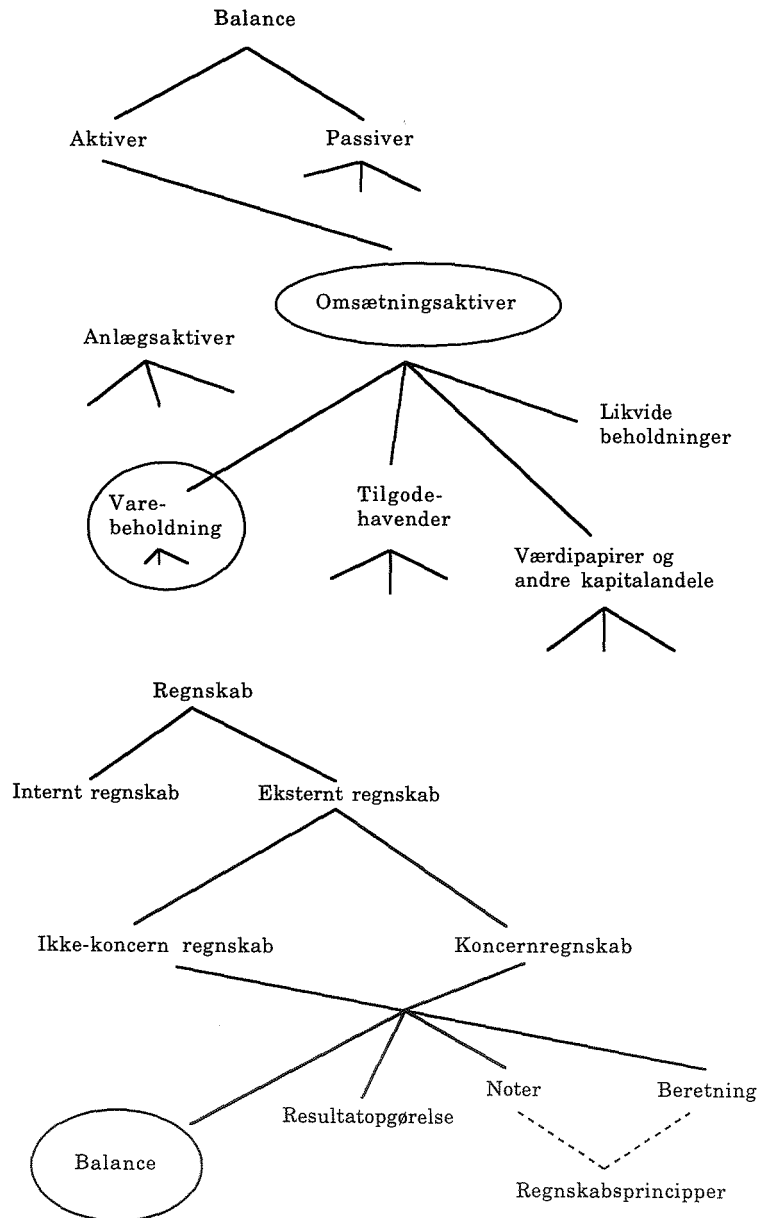
Det siger sig selv, at flerordsproblematikken og behovet for multispektion til opstilling af begrebssystematik og vurdering af synonymy er også til stede ved en automatisk sortering.

2.3 Udnyttelsen af tekstkorpus i lemmaselektionen

Vi mener på trods af ovennævnte skepsis omkring metoderne til lemmaselektion via et tekstkorpus fortsat, at et specielt sammensat tekstkorpus er det mest systematiske empiriske materiale og dermed bedste grundlag for udarbejdelsen af en systematisk fagordbog, og vi vil derfor give et bud på, hvorledes man også i selektionsfasen kan få en optimal udnyttelse af tekstkorpus. Vi vil give et par eksempler herpå fra vort regnskabskorpus, hvis danske delkorpus er på ca. 800 000 løbende ord.

I stedet for at anvende de to ekstreme muligheder, altså gennemlæsning af hele korpus og automatisk udskrivning af ordlister, har vi forsøgt at lade det tematiske/systematiske princip, som vi vil basere vores ordbog på, være styrende allerede i lemmaselektionsfasen. Vi har a

priori opdelt det univers, vi vil behandle, *det eksterne regnskab*, i en række deluniverser, se figur 1.



Figur 1.

Vi vil dernæst lade det indscannede tekstkorpus vise vej til de begreber, der indgår i de enkelte deluniverser. Vi vil illustrere denne fremgangsmåde ved at tage nogle eksempler fra deluniverset *omsætningsaktiver*. Vi søger ved hjælp af programmet *free-text-browser* på overbegrebet *omsætningsaktiver*. Begrebet indgår i årsregnskabet som en del af balancen, og da den danske årsregnskabslov regulerer opstillingen heraf, dvs. gruppering i romertalsposter (= de overordnede poster) og arabertalsposter (= underposterne) får vi herved foræret en struktur på dette delunivers. En række eksempler fra vort korpus viser, at terminologien her er meget ensartet, idet lovgivningens terminologi går igen i såvel lærebøger som autentiske regnskaber.

OMSÆTNINGSAKTIVER

- I Varebeholdninger
- 1 Råvarer og hjælpematerialer
- 2 Varer under fremstilling
- 3 Fremstillede færdigvarer og handelsvarer
- 4 Igangværende arbejder for fremmed regning
- 5 Forudbetaling for varer

OMSÆTNINGSAKTIVER

VAREBEHOLDNINGER

Råvarer og hjælpematerialer

Varer under fremstilling

Færdig- og handelsvarer

Igangværende arbejder for fremmed regning

Acontofakturering af igangværende arbejder

Forudbetalinger for varer

I ALT

Vi sætter os for at undersøge, om terminologien uden for opstillingerne er lige så konsekvent. Vi søger på *varelager*, som vi formoder er synonymt med de i opstillingerne anvendte *varebeholdninger* og finder et eksempel, der understøtter vor formodning:

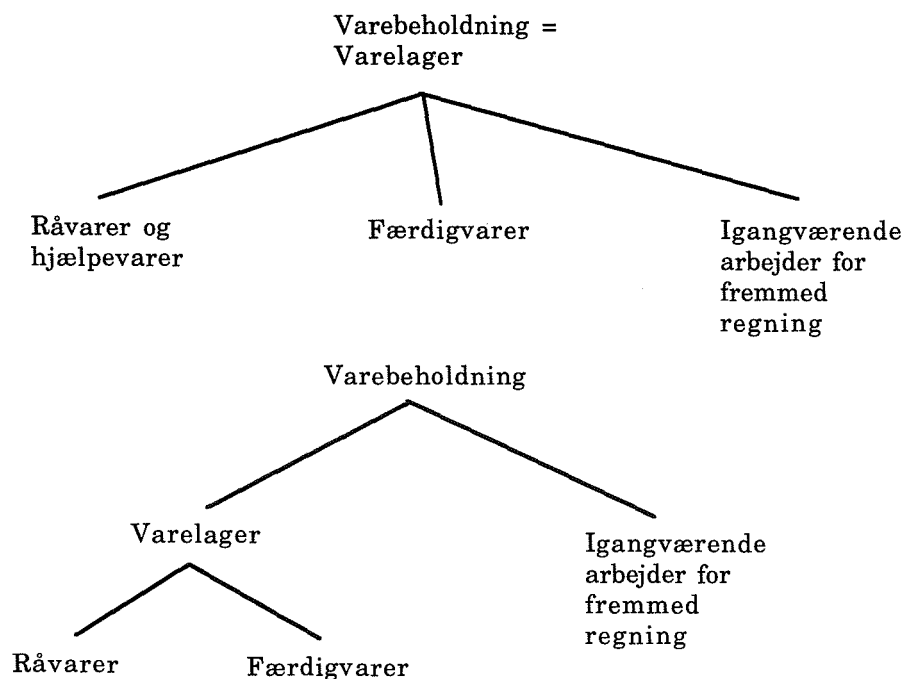
Varebeholdninger

Varelagre er optaget til laveste værdi af anskaffelses- eller genanskaffelsespris med nedskrivning for ukurans og langsom omsætningshastighed. (Noter — Autentisk regnskab)

Andre eksempler peger derimod i retning af, at dette ikke er tilfældet.

Når resultatopgørelsen således er påvirket af f.eks. afskrivningspolitik med hensyn til anlægsaktiver og værdiansættelsesmetoder for *varelager og igangværende arbejder*, vil valget af regnskabsprincipper på disse områder også påvirke balancen. (Lærebog)

hvor *varelager* og *igangværende arbejder* er sideordnede og dermed underbegreber til varebeholdninger. Vi har således to minibegrebssystemer som det vises på figur 2:



Figur 2.

Vi søger dernæst på nogle typiske kollokater til *varebeholdninger/varelager* (nedskrivning/værdiansættelse) og finder, at begge termer forekommer i denne kontekst, og ydermere dukker en tredje term op, som vi formoder er et synonym til varelager, nemlig *lagerbeholdninger*.

Det drejer sig om *nedskrivning på varelager*, henlæggelse til investeringsfond samt forskellen mellem skattemæssige og regnskabsmæssige afskrivninger på driftsmidler mv. (Lærebog)

Eventualskatten er beregnet som 40% af skattemæssige merafskrivninger på materielle anlægsaktiver og skattemæssige nedskrivninger på *varebeholdninger*. (Autentisk regnskab)

Når dækningsbidragstankegangen derfor anvendes ved *værdiansættelse af varebeholdninger* — dvs. uden tillæg af IPO — er der en inkonsekvens ved udarbejden af resultatopgørelsen. (Lærebog)

LIFO er ikke en alment accepteret metode til *værdiansættelse af lagerbeholdninger*, da den hverken viser den sædvanlig praksis eller giver de præcise resultatomkostninger. (Lærebog)

Figur 18.: Diagram over *værdiansættelse af varelager*. (Lærebog)

Vi har således en mistanke om, at der er forvirring omkring begrebsindholdet af disse termer og beslutter at gemme dem til eksperten (multispektion), idet de eksisterende

definitionsordbøger ikke kan klare spørgsmålet for os. Vi mener under alle omstændigheder, at alle tre termer skal medtages i lemmalisten, idet de, uanset frekvens, hører til begrebssystemet.

Vi vil vise endnu et eksempel på, hvordan vi via søgning på en overordnet term har fundet frem til en række termer, der kan bruges til tentativt at opstille en begrebssystematik i relation til det pågældende delunivers. Vi har valgt termen *afskrivning* og har ved søgning herpå fundet følgende flerordstermer:

Hvis der i stedet anvendes mere komplicerede afskrivningsregler, f.eks. *saldoafskrivning* i de første år og herefter *lineær afskrivning*, må redegørelsen være mere detaljeret. (Lærebog)

Eventualskat beregnes med 34 pct. af forskellen mellem den regnskabsmæssige og skattemæssige værdi af varelageret, omsætningsaktiver og hensættelser mv. samt af forskellen mellem de *regnskabsmæssige og skattemæssige afskrivninger* på materielle anlægsaktiver. (Autentisk regnskab)

Den skat, der kan henføres til årets resultat, er endvidere påvirket af varierende skattesatser i de enkelte lande samt af forskellen mellem *skattemæssige og driftsmæssige afskrivninger*. (Autentisk regnskab)

Akkumulerede afskrivninger 1. juni 1990

Kursregulering primo

Årets afskrivninger

Tilbageførte afskrivninger på solgte aktiver

Akk. afskrivninger 31. maj 1991

Bogført værdi 31. maj 1991 (Autentisk regnskab)

I praksis er denne regel kun aktuel for såkaldt *genvundne afskrivninger på bygninger mv.* (Lærebog)

Afskrivninger primo

Afskrivninger

Afgang

Afskrivninger ultimo (Autentisk regnskab)

Ekstraordinære afskrivninger (nedskrivninger)

De forudgående bemærkninger har taget sigte på de *ordinære afskrivninger*, dvs. afskrivninger, som foretages regelmæssigt (systematisk) år efter år. (Lærebog)

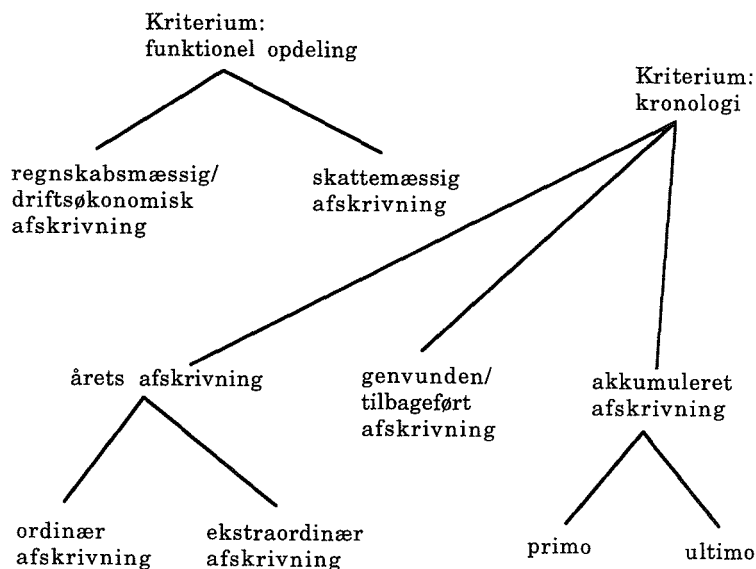
På basis heraf vil vi opstille et tentativt begrebssystem, som vises på figur 3.

Vi vil herefter lade en ekspert se på, om vores systematisering holder, herunder også bekræfte/afkræfte vores formodning om synonomi med hensyn til termerne *driftsmæssig/regnskabsmæssig afskrivning* og *genvundne/tilbageførte afskrivninger*.

Ved søgningen i tekstkorpus støder man i konteksterne til de pågældende termer ofte på de processer, der knytter sig til deluniverset *omsætningsaktiver*. For det første er der spørgsmålet, om selve **aktivering** af et beløb, dvs. om et beløb skal opføres som et aktiv (= **aktiveres**) eller ej. I tilfælde af aktivering skal genstanden/beløbet, etc. **opføres** i balancen, **værdiansættes** og beløbet kan senere **værdireguleres**. Sådanne termer tilknyttet

fordi disse er termer med en specifik betydning i forhold til andre universer (aktivering, optagelse, ansættelse), og dels fordi de refererer til generelle processer i forbindelse med de termer, vi fokuserer på. Ydermere har vi erfaring for, at især verberne (*aktivere, optage, ansætte*) volder problemer ved oversættelse, og vi vil derfor optage disse verbalformer i lemmalisten, ikke mindst på grund af deres høje frekvens i forhold til substantivformerne.

AFSKRIVNING



Figur 3. Tentativt begrebssystem

Andre værdipapirer og kapitalandele

Under denne post *opføres* f.eks. børsnoterede obligationer og aktier

Verbalformer: 268

Substantivformer: 0

Råvarer, hjælpematerialer og handelsvarer *optages til* anskaffelsespris eller dagspris, hvor denne er lavere.

For GN Store Nord er den væsentligste virkning af ændringer i årsregnskabsloven *optagelsen af* associerede virksomheder efter indre værdis metode

Verbalformer: 144 forekomster

Substantivformer: 1 forekomst

Råvarer og hjælpematerialer samt handelsvarer er *værdiansat til* indkøbspriser eller dagspriser, hvor disse er lavere.

LIFO er ikke en alment accepteret metode til *værdiansættelse* af lagerbeholdninger

Verbalformer: 130

Substantivformer: 185

3 Konklusion

Konkluderende kan vi sige, at undersøgelsen har vist, at et maskinlæsbart tekstkorpus *kan* være et nyttigt og hensigtsmæssigt redskab i forbindelse med selektion af lemmata til en specialordbog inden for fagsprog.

Korpus kan således *støtte*, dvs. af- og bekræfte leksikografens viden, og det kan *supplere* dennes viden (fx med hensyn til at finde — ukendte — underbegreber og synonymer), og desuden er det et hurtigt og effektivt redskab. Men en hensigtsmæssig anvendelse af korpus forudsætter bl.a. (1) at korpus er eksemplarisk, dvs. begrebsmæssigt dækkende sammensat og (2) at den leksikograf, der anvender korpus, har stor fagsproglig og stor faglig viden samt stor indsigt i terminologiske principper.

Litteratur

Ordbøger

Fryd 1990 = Ejnar Fryd: *Anglo/Amerikansk-Dansk, Dansk-Anglo/Amerikansk specialordbog inden for revision, regnskabsvæsen mv.* Tønder: FSRs Forlag.

Galindo 1989 = Orlando Galindo: *Erhvervs- og økonomiordbog Dansk-Spansk.* København: GTO.

Anden Litteratur

Bergenholtz, Henning/Jette Pedersen 1994: Zusammensetzung von Textkorpora für die Fachlexikographie. I: Burkhard Schaefer/Henning Bergenholtz (Hrsg.): *Fachlexikographie Fachwissen und seine Repräsentation in Wörterbüchern.* Tübingen: Gunter Narr Verlag, 161–176.

Bergenholtz, Henning/Sven Tarp m.fl. 1994: *Manual i fagleksikografi, Udarbejdelse af fagordbøger — problemer og løsningsforslag.* Herning: Systime.

Christiansen, Lisbet Maidahl/Grete Duvå/Anna-Lise Laursen 1994: Das Translationswörterbuch für Fachsprachen: Ein integriertes Konzept I: Burkhard Schaefer/Henning Bergenholtz (Hrsg.): *Fachlexikographie Fachwissen und seine Repräsentation in Wörterbüchern.* Tübingen: Gunter Narr Verlag, 269–282.

Duvå, Grete/Anna-Lise Laursen/Lisbet Maidahl 1992: Brugerundersøgelse vedrørende oversættelse af fagtekst. I: Ruth Vatvedt Fjeld (udg.): *Nordiske studier i leksikografi, rapport fra Konferanse om leksikografi i Norden 28.–31. mai 1991.* Oslo: Nordisk forening for leksikografi, 105–133.

Manual = Bergenholtz/Tarp 1994.

Pedersen, Jette 1995: *Koncernspecifik brancheordbog med særligt henblik på teknisk engelsk. Prækonceptionelle overvejelser og koncept.* Århus: Handelshøjskolen i Århus.