

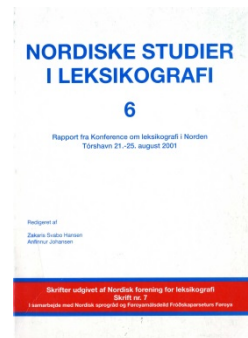
NORDISKE STUDIER I LEKSIKOGRAFI

Titel: Erfaringer fra arbejdet med Den Danske Ordbog

Forfatter: Ebba Hjorth

Kilde: Nordiske Studier i Leksikografi 6, 2003, s. 127-140
Rapport fra Konference om leksikografi i Norden, Tórshavn 21.-25. august 2001

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.

Erfaringer fra arbejdet med Den Danske Ordbog

Forbemærkninger

Jeg skal forudskikke at de erfaringer som jeg i det følgende gerne vil dele ud af, vil være sammensat af store linjer og overordnede betragtninger og detaljer i de enkelte ordbogsartikler.

Baggrund

Jeg vil dernæst – helt kortfattet – fortælle lidt om forudsætningerne for arbejdet med DDO: DDO bliver en stor ordbog i 6 bind over det moderne danske sprog. Ordbogen skal dække perioden fra 1950 og frem til udgivelsestidspunktet, som er 2002-2003. Arbejdet udføres i det danske Sprog- og Litteraturselskab, og ordbogen følger sig ind i rækken af de store videnskabelige ordbøger som Selskabet siden sin grundlæggelse har udgivet, fx Ordbog over det danske Sprog i 28 bind (udgivet 1919-1956), Holbergordbogen i fem bind (udgivet 1981-1988) og Supplement til Ordbog over det danske Sprog, som er planlagt til fem bind, hvoraf de fire er udkommet 1992-2001, og det femte er under udarbejdelse.

Plan for arbejdet

Arbejdet på DDO blev sat i værk på baggrund af en politisk vilje til og en beslutning om at gøre en indsats for det danske sprog i disse europæiserede og globaliserede tider.

- Der blev i 1989 holdt møder med deltagelse af interesserede institutioner og enkeltpersoner.
- En arbejdsgruppe blev nedsat til at udarbejde en plan for ordbogen.
- Det udarbejdede planforslag blev forelagt de oprindelige mødedeltagere.
- Planen blev færdigskrevet.
- Derefter begyndte forhandlingerne om finansiering af arbejdet med ordbogen.

Ordbogen endte med at blive bekostet af det danske kulturministerium og Carlsbergfondet i fællesskab med halvdelen til hver.

Planen indeholdt en beskrivelse af ordbogens anlæg, af dens ordforråd, dens metode, dens sprogsyn, dens empiriske grundlag, dens tidsforbrug og dens budget,

Det var nyt i den danske ordbogsverden at et ordbogsprojekt blev iværksat på baggrund af en så detaljeret plan.

Positivt – negativt

Og hvad var så positivt, og hvad var negativt ved at søge penge med udgangspunkt i denne detaljerede plan?

Positivt og – efter vores vurdering også nødvendigt – var det at have en plan med et fast budget og en fast tidsplan. Vi havde ganske enkelt ikke fået lov til at gå i gang med et projekt uden.

Negativt var det at vi i vores iver for at komme i gang med projektet og i vores stærke

tro på planens gennemførlighed ikke havde planlagt arbejdet som en totrinsraket, altså først havde søgt penge til et pilotprojekt hvor alle faser af arbejdet blev gennemprøvet. Og derefter på baggrund af erfaringerne fra pilotprojektet havde kunnet opstille realistiske tidsberegninger og budgetter som kunne bevilges i anden omgang.

Erfaringen er altså at ganske vist skal man vurdere hvor stort et beløb det er politisk muligt at søge til gennemførelse af et bestemt projekt, men det er u hensigtsmæssigt – af frygt for ikke at få pengene – at fastsætte beløbet uden et fuldstændig realistisk pilotprojekt, og det er u hensigtsmæssigt at sætte dette beløb for lavt. Jeg vender senere tilbage til hvad et sådant budget efter min mening skal tage højde for.

Ordbogens principper

Fra selve det overhovedet at have en fuld færdig plan før man begynder på et ordbogsprojekt, går jeg nu over til at beskrive enkeltelementer i planen.

I ordbogsplanen – som jeg i parentes bemærket selv har været med til at udarbejde – står der bl.a. følgende om principperne for ordbogen:

1. Ordbogen skal dække det nutidige danske sprog til og med udgivelsestidspunktet.
2. Ordbogen skal være en almensproglig ordbog der også medtager de alment brugte fagord.
3. Ordbogen skal udarbejdes på grundlag af et 40 millioner ord stort elektronisk tekstkorpus.
4. Ordbogen skal dække det skrevne sprog og inddrage det talte.
5. Ordbogen er deskriptiv i sit sigte. Den skal dog være normativ for oplysninger om ortografi og vejledende med hensyn til oplysninger om bøjning, ord dannelse, faste vendinger og konstruktionsmuligheder.
6. Ordbogen skal omfatte ca. 150.000 opslagsord.

Alle disse principper indeholder hver sine praktiske leksikografiske problemer som ikke nødvendigvis altid er forudset i en nok så smuk og gennemtænkt teoretisk plan. Hvilke af principperne der er de mest problematiske, kan ikke endeligt afgøres før ordbogen har stået sin prøve hos ordbogsbrugerne, men jeg vil i det følgende især beskæftige mig med princip 1, 3, 4 og 5.

Korpusset

Vi begynder med det tredje princip:

- Ordbogen skal udarbejdes på grundlag af et 40 millioner ord stort elektronisk tekstkorpus

At udarbejde en ordbog på grundlag af et så stort elektronisk tekstkorpus var nyt i 1989 da planen blev udarbejdet. Det var også nyt da arbejdet gik i gang i 1991.

Arbejdet med at etablere og nyttiggøre det elektroniske tekstkorpus havde tre problemkomplekser:

1. Hvordan skaffer vi 40 millioner elektroniske ord?

2. Hvordan klassificerer og strukturerer vi teksterne?
3. Hvilket programmel skal vi anvende?

Hvordan skaffede vi teksterne?

I begyndelsen var vi ansat fire ved projektet, og vi begyndte ved os selv idet vi groft udnyttede alle de forbindelser vi hver især havde til mulige tekstleverandører, og vi flød på den bølge af energi som et stort nyt og også nyskabende projekt giver.

Et par eksempler:

Iver Kjær, som er direktør for Det Danske Sprog- og Litteraturselskab, har hæk fælles med den daværende formand for folketingets præsidium. Over hækken fik Iver Kjær aftalt at vi kunne få stillet de første uredigerede afskrifter af forhandlingerne i det danske folketing til rådighed for ordbogsarbejdet. Vi kunne få så mange folketingsdebatter i WP-filer som vi kunne ønske os.

Min medredaktør, Kjeld Kristensen, er svoger til edb-chefen for et stort dansk bladhus. Svogeren var god at forhandle med og udvirkede at vi fik stillet maskinlæsbare tekster fra bladhusets aviser gratis til rådighed. Han forhandlede med de faglige organisationer i bladhuset og sørgede for at vi fik hjælp til konvertering af tekster til et brugeligt format.

Jeg havde selv nære forbindelser til en souschef i Danmarks Radio. Han formidlede kontakt til DR's generaldirektør, som blev interesseret i projektet og gav tilladelse til at alle afdelinger i DR måtte bruge tid på at finde udsendelser frem til os, være os behjælpelige med kopiering af video- og lydbånd og med at konvertere undertekster fra tekstede tv-udsendelser til et brugeligt format. Vi fik også mulighed for at leje rutinerede afskrivere i DR, som hjalp os med at afskrive de mange tv- og radioudsendelser.

En god ven og kollega, Henrik Holmberg, havde optaget og afskrevet annonceringer af togforsinkelser på københavnske S-togs-stationer. Han skulle bruge disse optagelser i forbindelse med et undervisningsforløb hvor han lærte DSB-ansatte at blive bedre til at meddele sig i perronhøjtalerne til de ventende S-togs-passagerer. Disse maskinlæsbare tekster blev stillet til rådighed for Den Danske Ordbogs korpus.

Gyldendals forlag, som til sin tid skal trykke ordbogen, var behjælpelig med at fremskaffe fotosatsfiler fra alle mulige af forlagets udgivelser. Og også de forlag som ikke vandt udgivelsesretten til ordbogen ved licitationen, var behjælpelige med at tilvejebringe tekster i elektronisk form og med at skaffe tilladelse fra de enkelte forfattere til at deres tekster kunne indgå i DDO's korpus.

Hver især hjemsøgte vi familiemedlemmer, tidligere og nuværende kolleger og bad om tekster. Vi ransagede egne gemmer for skilsmissepapirer, lejekontrakter, dagbøger, kæresterbreve, lejlighedssange og meget, meget mere.

Disse få eksempler blot for at give et lille indtryk af hvor vidt vi kom omkring. Men hvordan bar vi os ad med at få så mange mennesker til at rykke ud med deres tekster? Der var

nok flere grunde til at det gik så glat. For det første er Det Danske Sprog- og Litteraturselskab en offentlig kulturinstitution som er uden kommercielle interesser, og som nyder stor respekt i mange kredse. For det andet syntes de allerfleste som vi henvendte os til, at Den Danske Ordbog var et godt projekt som de gerne ville hjælpe på vej. For det tredje var der på det tidspunkt – og vi taler om årene 1991 og 1992 – ikke så stor frygt for hvad sådanne maskinlæsbare tekster eventuelt kunne bruges eller misbruges til, og endelig – og det er vigtigt – lovede vi hver enkelt tekstleverandør at deres tekster kun ville blive brugt til udarbejdelsen af DDO og til offentlig, ikke-kommerciel forskning. Dette sidste vender jeg tilbage til om lidt.

Positivt – negativt

Positive erfaringer var den interesse der var for at bidrage med empirisk materiale til en stor modersmålsordbog. Vi havde gode erfaringer med at bruge tid på at forklare mulige tekstleverandører hvorfor det var vigtigt at de bidrog med tekster, og på at forklare hvordan arbejdet ville komme til at foregå, og hvordan resultatet ville komme til at se ud.

Hvad der viste sig negativt, var at vi lovede tekstleverandørerne for meget. Det havde i mange situationer ikke været nødvendigt at aftale at de tekster som folk leverede, kun ville blive brugt til udarbejdelse af DDO. De restriktioner der er på brugen af DDO-korpusset, har givet os mange problemer. Det har været svært for mange at acceptere at DDO-korpusset kun kunne bruges på DDO's redaktion og kun til ikke-kommerciel forskning. Vi på ordbogen skal være de første til at beklage at vi ikke ved etableringen af korpusset kunne se hvor godt et værktøj det er til løsning af mange forskelligartede forskningsopgaver, og ikke har været i stand til at stille korpusset til rådighed i den grad vi gerne har villet, og i den grad som en ressource der er tilvejebragt for offentlige midler, har krav på.

Hvordan klassificerede og strukturerede vi teksterne?

Vi endte med at have 45.000 tekstenheder, og det var vores ambition at tekstkorpusset skulle være velbeskrevet og veldokumenteret. Korpusenhederne blev struktureret vha. SGML-standarden, og til hver eneste af de 45.000 tekstenheder blev der knyttet headeroplysninger til karakteristisk af den enkelte tekst. I headeren stod der disse oplysninger:

```
<Korpusenhed>
<Header>
<TxtOpl>

<Id>BfUe</Id>

<Restr>
<Ano>-</Ano>
<DDO>-</DDO>
</Restr>

<Ttit>JA TIL TYRKISK EVENTYR</Ttit>
<Vtit>B.T.</Vtit>
<Forl>-</Forl>
```

<Dat>
 <Dg>19</Dg>
 <Md>6</Md>
 <År>91</År>
 <Si>-</Si>
 </Dat>
 <Lo>1:31</Lo>

<AlFa>a</AlFa>
 <SkTa>s</SkTa>
 <RePr>r</RePr>
 <Arel>vv</Arel>
 <Medi>da</Medi>
 <Genr>?</Genr>
 <GnTy>?</GnTy>
 <Emne>795</Emne>

<Grp>BT910619.ONS</Grp>
 <Num>8420</Num>
 <Fil>BERL1E2</Fil>
 <Omf>613</Omf>
 </TxtOpl>

<SpbOpl>
 <EfN>Uno</EfN>
 <FoN>Steen</FoN>
 <Køn>m</Køn>
 <FøÅr>
 <År>1955</År>
 <Si>-</Si>
 </FøÅr>
 <FøS>Kbh.: Frb.</FøS>
 <Bop>?</Bop>
 <Reg>Ho</Reg>
 <Udd>journalist</Udd>
 <Erh>journalist</Erh>
 <SpV>i</SpV>
 <Rol>?</Rol>
 </SpbOpl>

Jeg skal ikke gennemgå disse oplysninger i detaljer, men blot pege på at teksten er dateret, den er klassificeret med hensyn til almensprog/fagsprog, talesprog/skriftsprog, reception/produktion, kommunikationssituation, medium, genre og emne. Desuden er sprogbrugens fulde navn, køn, alder, fødested, bosted, uddannelse og erhverv oplyst, og det er noteret om sproget kan karakteriseres som rigssprog eller som regionalsprog.

Man behøver ikke at se længe på disse headeroplysninger for at konstatere at beskrivelsen af de enkelte tekststykker er meget ambitiøs. Vi har brugt mange ressourcer på at opspore oplysningerne og på at fastlægge inventaret i de enkelte oplysningstyper.

Resultatet er at Den Danske Ordbogs korpus er velbeskrevet og velstruktureret i et alment format, og det er positivt at vi gjorde det. Det har givet os mulighed for mange fine restriktionsbeskrivelser rundt om i ordbogsartiklerne. Men erfaringen fra det daglige ordbogsarbejde viser at en del af oplysningerne ekstraheret fra headeroplysningerne kun gengives i ordbogsartiklerne i så ringe et omfang at anstrengelserne med at fremskaffe headeroplysningerne slet ikke står mål med de synlige resultater i ordbogen, og det er en negativ erfaring. Det gælder fx oplysningen om sprogbrugerens køn. Det har ikke været muligt at finde en artikel i ordbogsmanuskriptet der beskriver et ord, en betydning eller en konstruktion som især tilhører kvinders eller især tilhører mænds sprogbrug. Det nærmeste jeg har kunnet komme noget sådant, er artiklen *søstersolidaritet*, der i øjeblikket ser således ud:

søstersolidaritet sb.

fk. -en

1 kvinders solidaritet med hinanden især brugt blandt medlemmer af kvindebevægelsen i 1970'erne.

2 søstres solidaritet med hinanden (sj.).

HIST: 1970

Men oplysningerne er i mange situationer udnyttet til andre formål end det egentlige ordbogsarbejde, og vi kan kun håbe på at det sker i endnu højere grad i fremtidens brug af DDO's korpus.

Konkluderende kan man sige at DDO's korpus har vist sig som et særdeles fremragende grundlag for redaktionsarbejdet. Som man kan forstå, har vi gjort meget ud af at sammensætte korpusset meget bredt. Det indeholder både skriftsprog og talesprog, både almensprog og alment brugt fagsprog, sprog produceret til mange mennesker og sprog produceret i private sammenhænge. Mange genrer er repræsenteret, og sprogbrugernes alder strækker sig fra helt unge mennesker til gamle. Tidsmæssigt er teksterne i DDO's korpus fra 1983 til 1992, den periode som vi kalder ordbogens primærperiode. Da vi udformede planen, mente vi at afstanden mellem primærperioden og ordbogens udgivelsestidspunkt var tilpas lille til at vi kunne hævde at vi dækkede også det allernyeste sprog. Imidlertid måtte vi på et tidspunkt ud i nye forhandlinger med Carlsbergfondet og Kulturministeriet om tilførsel af flere midler, således som jeg gjorde rede for det ved den leksikografiske konference på Hanaholmen. Det betød at ordbogens udgivelsestidspunkt fjernede sig 3-4 år. Dette har givet os anledning til at vi nu også i redaktionsarbejdet er begyndt at anvende andre korpora for at kunne dække også de allerseneeste år, og for at citaterne i ordbogen ikke alle sammen daterer sig til det gamle årtusinde. Denne praksis kan naturligvis diskuteres. Og det bliver den på redaktionen. For det første kan man sige at det giver en skævhed i beskrivelsen at en del af artiklerne bygger på DDO-korpusset, mens andre inddrager nyere materiale fra korpora og fra nettet, som redaktørerne nu har adgang til. Et andet moment i denne praksis er tidssforbruget. Det tager tid at søge på nettet efter et bestemt ord eller en bestemt ordforbindelse. Og tid er ikke det vi har mest

af. Så på et tidspunkt var det nødvendigt at diskutere følgende: Har vi med kun 7 måneder tilbage til at færdiggøre manuskriptet i ressourcer til dette arbejde? Nogle redaktører syntes det tog for lang tid, og at det gik ud over akkorden at bruge andre korpora og nettet. Andre syntes at netsøgning i nogle tilfælde var hurtigere end søgning i vores eget korpus. Så den enkelte redaktør fandt sin egen praksis under hensyntagen først og fremmest til akkorden. Sådan var vi nødt til at tænke.

Men på spørgsmålet "Har Den Danske Ordbogs korpus vist sig at have en passende størrelse?" må svaret være både ja og nej. I mange tilfælde – de fleste – er korpusset passende stort til at alle aspekter ved et ord kan fremfindes i korpusset. I andre tilfælde er det for lille og må suppleres med andet materiale. I atter andre tilfælde er forekomsterne af et bestemt ord eller fænomen alt for mange til at materialet kan overskues og gennemarbejdes. I sådanne tilfælde har vi måttet reducere i materialet.

Norm – usus

Nu går vi videre til det femte princip:

- Ordbogen er deskriptiv i sit sigte. Den skal dog være normativ for oplysninger om ortografi og vejledende med hensyn til oplysninger om bøjning, ord dannelse, faste vendinger og konstruktionsmuligheder.

Denne passus i planen for ordbogen er nok den der har givet anledning til størst hovedbrud og til de fleste diskussioner på redaktionen.

Ortografi

Med hensyn til opslagsordenes ortografi – som vi altså if. vores kommissorium skal angive normen for – har vi valgt også at beskrive andre staveformer end de officielle, altså dem der står i Retskrivningsordbogen udgivet af Dansk Sprognævn.

Et par eksempler:

Grundstoffet *klor* staves iflg RO k-l-o-r. Det er den officielle form, og det er den vi if. vores plan skal bruge. Imidlertid har vi valgt at lade artiklen se således ud:

klor sb. (uofficiel stavemåde, alm. blandt kemikere: chlør)

fk. el. itk. -en el. -et

[,klo?R]

grundstof nr. 17 symbol: Cl; hører til halogenerne □ *freon spaltes i frit ustabil klor, som „æder“ ozonen* Jørgen Steen: Geografi i syvende Lærerens bog (1989)

- gulgrøn og giftig luftart med stikkende lugt; kemisk særdeles reaktionsdygtig □ *Forslagsstillerne henviser til .. udslippet af 150 kg klor i efteråret 1978, hvor 14 mennesker måtte indlægges til observation for klorforgiftning Status 1985, I dagligsproget bedder det „der er chlør i vandet“, selvom der egentlig er tale om chlørundersyring* Paul Jespersgaard: Kemi 1 (1988).

AFL: SUFF: →sb. klorat, klorid →vb. klore, klorere

SMS: kloratom, -forbindelse, -gas

HIST: af gr. chloros ‚lysegrøn‘

Vi har altså valgt at supplere den officielle norm med en angivelse af hvordan fagfolk staver til ordet hvis de kan komme af sted med det.

Morfologi

På samme måde med bøjningsoplysninger:

Også her angiver vi den officielle norm, men er der andre muligheder som er repræsenteret i vores korpus i tilstrækkelig mængde, anfører vi også dem, men tager vores pligt til at vejlede sprogbrowseren alvorligt ved at beskrive denne ikke officielle form som uofficiel.

Det ser fx således ud:

job sb.

itk. -bet, - el. (uofficielt) =s, -bene el. (uofficielt) =sene

[,d;Cb]

1 fysisk el. åndelig virksomhed som man til stadighed udøver for at forsørge sig selv (og sin familie) fx som ansat i en virksomhed = mere formelt arbejde; typisk: *civilt ~, krævende ~, nyt ~, spændende ~, hårdt ~, fast ~; bestride et ~, få tilbudt et ~, miste sit ~, søge ~, passe sit ~, ~ og uddannelse, ~ og karriere, ~ inden for det offentlige* □ *Et godt job med masser af penge ville da være dejligt, Hun var lykkelig for sit job som ekspedient i Magasin* • sted hvor man udøver denne virksomhed (kendt fra 1968) = arbejdsplads *uden pl.*; typisk: være på ~ □ *For os er det en stor forskel at arbejde samme sted, fordi vi nu kan køre sammen til og fra jobbet.*

2 opgave el. stykke arbejde der skal udføres fx på arbejdspladsen el. i hjemmet = arbejdsopgave □ *Siden jeg var ti år, har det været mit job at dele gaver ud (: til jul).*

AFL: til bet. 1 SUFF: →vb. **jobbe** →adj. **jobmæssig**

SMS: **jobfunktion**; til bet. 1 **jobannonce**, -marked, -rotation, -situation, -søgning; *bi-, deltids-, drømme-, fritids-, fuldtids-, halvdags-, kvinde-, lærer-, mande-, rengørings-, sekretær-, *skåne-, vikar-, ønskejob

Substantivet *job* hedder iflg. RO *job* i ubestemt flertal og *jobbene* i bestemt flertal. Den ubestemte pluralisform *jobs* er imidlertid kolossalt frekvent i DDO-korpusset og ubestemt pluralis *jobsene* forekommer. Derfor har vi valgt at indrette bøjningsdelen som vist ovenfor. Vi angiver normen, men beskriver også brugen, og vi vejleder ordbogsbrowseren ved at klasificere *jobs* og *jobsene* som uofficielle.

Semantik

En lidt anden – og måske lidt mere diskutabel – form for normering og vejledning anvender vi ved betydningsbeskrivelser:

Et eksempel er ordet *betænksom*:

Det kommer i DDO til at se således ud:

betænksom adj.

-t, -me

1 som tænker på og tager hensyn til andre om personer jf. hensynsfuld □ *det var sødt og betænksomt af hendes far at give ham det arbejde*

• som vidner om tanke på andre □ *(det falder) naturligt at pege på et medlemskab af Det kgl. Haveselskab som en betænksom julegave.*

2 (denne brug regnes af mange for ukorrekt) = "betænkelig □ *Der er dog flere ting, som gør aktieeksperterne betænksomme ved direkte at anbefale investorerne at købe aktien på nuværende kursniveau.*

AFL: til bet. 1 SUFF: →sb. betænksomhed

Vi kan i vores korpus konstatere at betydning 2 forekommer. Derfor beskriver vi den sådan som vi skal if. vores plan, men vi har valgt at give den et par vejledende ord med på vejen, nemlig disse *denne brug regnes af mange for ukorrekt*. På tilsvarende måde kan vi beskrive fx bestemte syntaktiske konstruktioner som er belagt i korpusset, men som ikke er helt 'stuerene'.

Vi betragter disse bemærkninger som en del af vores vejledende opgave, og det mener vi er positivt, men denne vurdering bygger jo i virkeligheden først og fremmest på redaktionens egen opfattelse og eventuelt på den sproglige debat sådan som den fx kommer til udtryk i læserbreve når visse sprogbrugere brokker sig over sædernes forfald. Og det er måske negativt.

Vi er meget spændte på hvordan denne balancegang mellem norm og usus vil blive modtaget af de kommende ordbogsbrugere. Nogle sprogbrugere glæder sig over at ord skifter betydning, og synes at det er udtryk for at sproget er levende; andre harmes over det og synes netop at det er udtryk for slaphed og, som nævnt ovenfor, sædernes forfald.

Talesprog

Det fjerde princip i kommissoriet var:

- Ordbogen skal dække det skrevne sprog og inddrage det talte

For at leve op til dette krav gjorde vi os – som jeg allerede har beskrevet det – stor umage for at få indsamlet talesprog til vores korpus.

Talesprogsteksterne kommer til udtryk i ordbogen inden for forskellige oplysningstyper:

For det første som ord der kun eller især forekommer i talesproget.

Et eksempel på dette er bandeord:

dæleme adv.

[,dE:l@m@]

bruges som mildt, forstærkende bandeord for at understrege en personlig holdning, forsikring m.m. (især talesprog) = søreme, knageme, sandelig □ *Så let skal hun dæleme ikke slip', den møkke!*

dæleme om ikke .. = "gudhjælpemig om ikke ..

HIST: sammentrækning af *djævlen* (tage e.l.) *mig*

Et andet eksempel er udråbsord:

ad¹ udråb.

[,áDʔ]

bruges for at udtrykke væmmelse (især talesprog) = føj, bvadr □ *Ad, se hvor han sidder der og mæsker sig med sin drink, det fede svin.*

HIST: vel lydefterlignende

Men også mere almindelige ord kan være begrænset til talesproget, fx

dersens adv.

[,dERʔs@ns] el. [,deRʔ-]

bruges for at lægge afstand til el. tage forbehold over for noget omtalt el. nærværende (især talesprog) = der, jf. hersens [efter personligt el. demonstrativt pron.] □ *så ham dersens blikkenslageren .. så kom han en dag ned på tankstationen og ville snakke, hvis de dersens muslimer ikke vil tilpasse sig, kan de jo bare blive hjemme.*

HIST: dannet af *der* efter *hersens*

Endnu et eksempel er:

-én el. **-en** sidsteled

fk.

genstand som man ikke kender betegnelsen på el. ikke lige har lyst til at betegne nærmere (især talesprog, uformelt) [med sb. som betegner et materiale, som førsteled, fx *pap-én*] □ *Den ene så ud til at være en vinflaske. Den anden var bare sådan en blå plastic-én, der havde været sprit i.*

Ord kan have særlige betydninger som kun findes i talesprog, og særlige talesprogs vendinger kommer med i DDO med baggrund i forekomsten i talesprogsteksterne, fx vendinger som *og alt sådan noget* og *og alt det dér*.

Også særlige syntaktiske konstruktioner er i sjældne tilfælde registreret. Et eksempel på dette er:

at², infinitivpartikel

....

8 bruges i visse konstruktioner direkte foran en retningsangivelse uden et verbum imellem (især talesprog) □ *Mon de da ikke snart skulle til at hjemad!, Så blev det tiden, jeg skulle til at i skole.*

Vi betragter det som noget meget positivt ved vores ordbog at talesprog er inddraget så meget som det er, men det har været besværligt og tidkrævende at arbejde med talesprogsteksterne. Det hænger sammen med at en stor del af det talesprogsmateriale der indgår i vores korpus, er meget fintmærkende skrevet ud fordi det er bragt til veje med andre formål end ordbogsarbejde. Det har gjort det besværligt at søge i, og specielt har det gjort det

vanskeligt at finde autentiske citater som på én gang er korte, selvforklarende og udgør en syntaktisk helhed. Følgende lille note fra en redaktør illustrerer meget godt problemerne med at bruge talesprogsteksterne:

Subsem
 #NOTE
 Til: Kom.
 Ang.:
 Fra: jhp
 Ann. som 4
 Semspec ofø
 Denbet sammensætte på en lidt heldig el. smart måde
 Genprox sammensætte
 Onym
 Ordfelt
 txt bikse sammen
 Valens
 Aktant NGN brygger NGT (sammen)
 Dok
 Citat
 #NOTE
 Til: Kom.
 Ang.:
 Fra: jhp
 Har redigeret lidt i talesproget
 txt *nå men så får man da brygget lidt mad sammen og så og hevet
 lidt vin frem og .. så hygger man sig*
 Kilde
 Kildeid HRNO

Den autentiske tekst ser således ud:

<replik id=ZZ1>jaja {pause} nå men så får man o- får man da brygget lidt mad sammen og så {pause} og hevet lidt vin frem og {tøven} s- nå ja men så hygger man sig

Medarbejdere

Hvordan er det lykkedes os at skaffe kvalificerede medarbejdere til så stort et projekt? Det er det først og fremmest fordi kvalificeringen er sket inden for projektets rammer. Men de medarbejdere som vi har villet satse på at uddanne til leksikografer, har jo skullet udvælges! Og da har det været vores erfaring at kandidater med dansk som modersmål med et fremmedsprog, med lingvistik og med datalingvistik er blevet dygtige leksikografer. Det lyder ejendommeligt at kravet om en kandidateksamen med dansk som fag slet ikke er inde i min vurdering, men vi har konstateret at de sproglige discipliner i danskstudiet ikke har været af en sådan art og af et sådant omfang at det har slået kandidater med den kontrastive vinkel på dansk med udgangspunkt i et fremmedsprog af banen.

De egenskaber vi har lagt vægt på er følgende:

- Evne til at analysere og strukturere store datamængder.
- Evne til at formulere en god betydningsbeskrivelse.
- Evne til at gøre tingene færdige.
- Evne til at tage en beslutning.
- Akribi
- Og frem for alt: kærlighed til et så nørdet arbejde som det er at skrive ordbogsartikler dag ud og dag ind.

Ved udvælgelsen af kandidater fik vi efterhånden glæde af at udlevere en mindre konkordans til mulige redaktører og bede dem ved en samtale forholde sig til konkordansen ud fra nogle oplysninger om hvilket projekt DDO er. Det gav et godt indblik i ansøgernes kvalifikationer til leksikografisk arbejde.

Det er min opfattelse at der er tre elementer i leksikografisk arbejde:

For det første en sproglig grundviden.

Dernæst et leksikografisk håndværk.

Disse to elementer kan læres, men det tredje element, som er kombinationen af de seks egenskaber som er nævnt ovenfor, kan ikke læres. Enten kan man det, eller også kan man det ikke!

Tidsbegrænsning

Som jeg nævnte tidligere, var vi selv med til at sætte rammerne for dette projekt. Ud fra en vurdering af hvad det var politisk muligt at opnå, fastsatte vi en økonomisk og tidsmæssig ramme for arbejdet. Disse vurderinger hvilede på for tyndt et grundlag, og vi måtte som sagt gå Canossagang til pengekasserne.

Vi fik en ny bevilling, og vi fik konstrueret et administrationssystem som – lidt forenklet sagt – kan holde rede på hvor mange ‘planlinjer’, som vi kalder de enkelte linjer i ordbogen, der skrives hver dag og af hvem, og hvor mange planlinjer der nu er tilbage at skrive. Det blev beregnet hvor meget vi skal nå hver dag, ved at regne baglæns fra afleveringstidspunktet efter at have talt op hvor mange planlinjer der skal skrives i alt, hvor mange ord der – cirka – skal skrives, og hvor mange linjer hvert enkelt ord må fylde.

Men det at have et administrationssystem gør det jo ikke alene. For hvad nu hvis redaktørerne ikke når at skrive eller gennemlæse de planlinjer som de skal? Eller hvad nu hvis redaktører skal have barselsorlov? Eller hvis de finder andet arbejde før manuskriptet er færdigt? Vi har været gode til at udregne planlinjetal og holde øje med udviklingen, men vi har ikke været gode til at sikre midler til at afhjælpe sådanne katastrofesituationer.

Vi har været gode til at gennemføre vores politik, som har været denne:

Vi skriver hver især hvad vi kan. Nogle skriver meget af høj kvalitet. Andre skriver mindre enten af høj eller af ringere kvalitet, men vi er solidariske med hinanden, vi trækker på samme hammel, og vi er fælles om projektet!

Men ja, der er da fordele ved at et ordbogsprojekt er tidsbegrænset:

Der er den fordel at der kommer en ordbog. Men ulempen ved tidsbegrænsningen er at ord-

bogen ikke bliver god nok. Hvis ordbogsprojektet ikke var tidsbegrænset, så ville ordbogen blive god, til gengæld ville den så aldrig blive færdig!

Tidsbegrænsning er en fordel hvis man selv har haft kvalificeret indflydelse på den. Men tidsbegrænsningen er en pestilens i den sidste del af perioden.

Tidsbegrænsning er en nødvendighed, men man må aldrig fastsætte tidsbegrænsninger for et ordbogsprojekt uden at indregne:

- Mulighed for efteruddannelse til medarbejderne.
- Mulighed for medarbejdernes private forskning.
- Tid til løbende formidling af erfaringer og resultater.
- Tid til strategisk, fremadrettet arbejde med nyudvikling af ordbøger, arbejdsmetoder og arbejdsredskaber.
- Tid til at skrive ansøgninger til nye projekter.
- Sygdom og barsler.

Det gjorde vi ikke i tilstrækkeligt omfang, men det gør vi næste gang!

Anfægtelser

Men anfægtelser, det får man undervejs når man arbejder så længe på så stort et projekt som DDO.

Anfægtelser der går på følgende spørgsmål:

Hvordan skal en ordbog se ud i fremtiden?

Hvilke oplysninger skal fortsat rummes i en traditionel, redigeret ordbogsartikel, og hvilke kan man mere effektivt på et større materiale finde i velbeskrevne korpora forsynet med velegnede søgesystemer?

Skal vi fortsat bruge krudt på at konstruere en elegant betydningsbeskrivelse der medtager alle nødvendige elementer af en bestemt ting, når vi bliver meget klogere af at se et billede, en filmstump eller andet der viser tingen?

Skal vi fortsat kæmpe med norm-beskrivelser i ordbøger hvis sprogbrugerne lige så let selv kan finde frem til oplysninger gennem søgninger i deklarerede korpora, og det hele alligevel kan være i mobiltelefonen?

