
Den Danske Ordbog - set i bakspejlet

*Af Henrik Lorentzen og
Lars Trap-Jensen*

Med udsendelsen af sjette og sidste bind i november 2005 fremstår Den Danske Ordbog nu som et færdigt værk – 14 år efter at redaktørerne gik i gang med arbejdet. Med værkets afslutning skifter det uundgåeligt karakter, ikke mindst for de redaktører der har levet i nær kontakt med projektet en god del af deres arbejdsliv. Fra at have været en foranderlig størrelse, genstand for diskussioner og programerklæringer, har ordbogen fået selvstændig status som værk, og nu er tiden kommet til at anskue den med den distance der følger deraf.

I den tid der er gået siden arbejdet med Den Danske Ordbog (herefter DDO) begyndte, er der blevet både sagt og skrevet en del om ordbogsarbejdet, også i dette blad. I en artikel fra 1999 (*Nyt fra Sprognævnet* nr. 3, 1999) redegjorde Ebba Hjorth for ordbogens forskellige opgaver med at normere, beskrive, vejlede og dokumentere sproget. Det vil vi ikke gentage her, men derimod koncentrere os om at beskrive de træk der i særlig grad karakteriserer

DDO og gør den til en anden slags ordbog end de øvrige ordbøger der findes på det danske marked.

En korpusbaseret ordbog

Skal der fremhæves et enkelt træk hvorved DDO adskiller sig fra andre danske ordbøger, er der ikke tvivl om at det er dens kildegrundlag. De første par år af projektets levetid gik med at designe opbygningen af et tekstkorpus og fremskaffe, formatere og opmærke de 44.000 forskellige tekstprøver med i alt 40 millioner løbende ord som skulle udgøre DDO's kildegrundlag. Og frem for noget er det det der gør DDO til en ny og anderledes ordbog, både med hensyn til arbejdsmetoder og -betingelser for dem der har stået for redigeringen af artiklerne, og for ordbogens brugere, der næppe kan undgå at bemærke det når de slår op i ordbogen. DDO er den første korpusbaserede ordbog der er udarbejdet for dansk, og det har sat sit præg på ordbogen i en grad som måske ikke er helt indlysende for den der blot slår nogle få ord op.

Noget af det der adskiller et korpus fra en stor tekstsamling som fx artikelbasen Infomedia eller alle de mange hjemmesider der kan søges fra en søgemaskine som Google, er at korpusteksterne er udvalgt med særligt henblik på at repræsentere en bred vifte af både teksttyper og sprogbrugere. DDO's korpus er ikke 100 % repræsentativt for alt sprog der produceres i Danmark – og mange korpuslingvister mener heller ikke det er muligt, måske ikke engang ønskeligt, at opnå fuldstændig repræsentativitet – men der er tale om et varieret

og bredt sammensat korpus, det man kalder et balanceret korpus. Et andet punkt der adskiller tekstkorpuset fra de rå tekstsamlinger, er at alle korpussets tekster er blevet opmærket med oplysninger om teksten (dens titel, dato for tilblivelsen, hvilken genre den tilhører, hvilket medium den er produceret til, hvilket fagligt område den tilhører, osv.) og om den eller de sprogbrugere der har produceret teksten (køn, alder, bopæl, kommunikationssituation osv.). Det gør det muligt at lave statistiske undersøgelser af sammenhængen mellem et sprogligt fænomen og en bestemt teksttype eller en bestemt sociologisk variabel. I ordbogen kan man se det i form af forskellige etiketter der angiver om fx en betydning især er hyppig i en bestemt teksttype eller hos en bestemt gruppe sprogbrugere. Der kan stå *især yngre*, *især talesprog*, *især i bibelske tekster*, *nu sjældent* og en række andre oplysninger. Disse angivelser er altså ikke kun fremkommet efter redaktørernes personlige skøn, men bygger på en kvantitativ analyse af tekstgrundlaget. Det kan også være oplysninger om sproglige former: Hvis materialet viser at en bestemt bøjningsform er dominerende, eller at en tilladt dobbeltform stort set ikke bruges, oplyses der om det med etiketter som *især i præet. ptc. brugt som adj.* eller *næsten kun i formen chefstol* (frem for *chefsstol*).

Derudover er der forskellige oplysningstyper hvor korpusgrundlaget ikke kan aflæses direkte, men hvor det ikke desto mindre har haft stor betydning for resultatet. Lad os se på de mest oplagte tilfælde.

Ord og betydninger

De ord der er taget med som opslagsord i DDO, er blevet udvalgt på grundlag af redaktionens kilder, og her har korpusset været den vigtigste. Det betyder at faktisk forekomst og hyppighed i sproget har været et vigtigt kriterium. Kun i begrænset omfang har vi medtaget ord som ikke var belagt i korpusmaterialet. Vi har villet rydde ud i de "ord-lig" som nemt kan forekomme når nye ordbøger overtager ordstof fra eksisterende ordbøger og glemmer at kassere de ord som er gået af brug. Af sådanne forældede ord kan nævnes *abildgård*, *drunte*, *dystløb*, *fjermer* og *fortrøste* – ord der står i Sprognævnets retskrivningsordbog, men som ikke er kommet med i DDO. Princippet om at lægge så stor vægt på faktiske forekomster viser både styrken og vel også den risiko der er forbundet med denne måde at arbejde på. En åbenlys styrke er at man derved får øje på hidtil upåagtede ord og betydninger fra kroge af sproget som normalt ikke er tilgængelige for en ordbogsforfatter. Det gælder fx det private sprog fra breve, dagbøger og lejlighedssange, og det gælder især talesproget. Næsten en femtedel af teksterne i ordbogens korpus udgøres af udskrevet talesprog, og det gør at talesproget har fået en fyldigere behandling i ordbogen end man ser det i andre ordbøger. Af opslagsord er det fx udråbsord og kraftudtryk som *ad*, *arh*, *dæleme*, *kraftstejleme*, *shit*, *sørenjenseme*, *uf*, *yes* og *æh* der især hører til talesproget, og i en artikel som *altså* kan man se hvordan det har sat sit præg på betydningsopdelingen: Fem ud af otte betydninger er markeret som talesprog. Desuden er

der særlige grammatiske konstruktioner som er typiske for talesproget; også det oplyses der om: 'Tag *en til kage*' (frem for *en kage til*), 'hun kommer *fra Fyn af*', 'det vil vi *slet, slet ikke*', 'jeg laver film *plus at* jeg er meget sammen med mine venner'.

Hvis vi tager de selvkritiske briller på, kan vi også godt få øje på begrænsninger i orddudvalget. Selvom forekomsttal ikke har været det eneste kriterium for udvælgelsen, har det spillet en stor rolle, undertiden også lidt for stor. Det har i hvert fald ført til at der er sluppet mange simple ordsammensætninger med, enten som opslagsord eller som eksempler på orddannelser uden egen definition, og det er sket på bekostning af ord som nok er sjældnere, men i de fleste tilfælde også sværere at forstå. *Gigtplaget*, *kageform* og *telefonregning* er eksempler på udbredte, men let genemskelige ord som har fået deres egen artikel, mens man slår forgæves op efter *bassethorn*, *kohorte* og *transformationsgrammatik*. Det er ikke fordi de sidstnævnte ord er blevet overset af redaktionen; de er blevet fravalgt fordi der har været for få forekomster i korpus til at de kunne optages efter ordbogens principper. Men der er nok dem der vil mene at det ofte er de sjældne eller faglige ord man har brug for at slå op i sin ordbog.

Konstruktioner

Det store elektroniske tekstkorpus har givet redaktørerne et solidt grundlag for at beskrive ikke bare hvordan danske ord staves og bøjes, og hvad de betyder, men også hvordan de sættes sammen til

sætninger. Derfor er der ved samtlige verber i ordbogen givet oplysninger om hvilke sætningsled de typisk knytter til sig. Disse oplysninger er sat med en særlig skrifttype og står i en firkantet parentes efter betydningsangivelsen og før sprogbrugseksemplerne.

glo vb.

-r, -ede, -et; ['glo']

se el. stirre med vidt åbne øjne;

[NGN glør (på/efter NGN/NGT), NGN glør

+RETNING] □ ~ olmt/surt, ~ måbende/

nysgerrigt □ han gloede, så øjnene var

ved at falde ud af hovedet på ham

DRTV-drama89

- kigge på en ukritisk el. passiv måde

(nedsættende); [NGN glør (NGT/på_NGT)]

□ ~ fjernsyn, sidde og ~ ud i luften

□ han sad .. samtlige seks timer bare

og gloede tegnefilm talespKbh87

AFL PRÆF → beglo vb.

HIST oldn. glóa, eng. glow 'gløde',

jf. gloende, glohed, glørød.

Konstruktionsoplysningerne har et formelt præg, men for at de skal være lette at forstå og selv fungere som mønstre, har redaktionen valgt en notation der ligger forholdsvis tæt på dagligsproget. Ved verbet *glo* finder man således sætningsmønstret [NGN glør (på/efter NGN/NGT)], som betyder at man roligt kan danne følgende sætninger: *Vi gloede på den gule bus* og *Vi gloede efter pigen*. Muligheden for ikke at angive hvad man glør på eller efter, er udnyttet i det anførte citat: *han gloede, så øjnene var ved at falde ud af hovedet på ham*. Derudover gives også mønstret [NGN glør +RETNING], der altså giver mulighed for

at danne sætninger med andre retningsangivelser som fx *han gloede ud ad vinduet* eller *hun glør hen på mig*.

Ved hjælp af konstruktionsoplysningerne kan man altså danne danske sætninger der er typiske og mundrette, i hvert fald hvad syntaksen angår. Mønstrene er nemlig udvalgt så de afspejler de konstruktioner redaktørerne typisk har set i korpus. Men hvad med mere kontroversielle sætningskonstruktioner? En konstruktion der ofte påkalder sig negativ opmærksomhed, er *kommentere på*. I artiklen *kommentere* er der taget højde for den, idet den traditionelle konstruktion [NGN kommenterer NGT/NGN] gives først efterfulgt af [NGN kommenterer på NGT], som ledsages af bemærkningen *konstruktionen med på regnes af nogle for ukorrekt*. På denne måde imødekommer ordbogen sit dobbelte formål, nemlig både at beskrive og vejlede, idet den forkætrede sprogbrug beskrives samtidig med at man gør opmærksom på at den risikerer at medføre løftede øjenbryn hos visse modtagere.

Oplysningerne om konstruktion er obligatoriske ved verberne i DDO, men desuden findes der ofte også i andre tilfælde, især ved substantiver og adjektiver, oplysninger om hvilke konstruktioner ordene indgår i. Her er oplysningen ikke obligatorisk, men er givet hvis en bestemt sprogbrug er særlig fremtrædende i korpus. Og her vil mange problemer ved tekstproduktion kunne løses, fx: Hvilken præposition bruger man typisk efter *jaloux – på, med, i* eller måske *over*? Ordbogen giver svaret: [jaloux på NGN] og [jaloux over at../NGT]. Tilsvarende får man i

artiklen *egenskab* oplyst at det typisk hedder en [egenskab ved NGN/NGT] eller [hos NGN/NGT]. Det er nyt at en dansk ordbog informerer i så stort omfang om disse forhold, der er meget vigtige for produktion af tekst, og det er nyt at oplysningerne gives i formaliseret, men alligevel – mener og håber vi – let tilgængelig form. Foreløbig har vi ikke modtaget så mange tilbagemeldinger om lige præcis denne oplysningstype, men vi håber at brugerne efterhånden vil få øjnene op for de store muligheder der ligger i at bruge ordbogens oplysninger om bl.a. valg af præposition.

Kollokationer

Betegnelsen dækker over ordforbindelser som ikke kræver nogen forklaring, men som gives som eksempler på almindelig sprogbrug (*sove trygt, blussende rød, stille et forslag, begå et røveri, skæg og ballade*). DDO bringer en række af den slags korte sprogrøver uden forklaring, og de er også fremkommet efter undersøgelser i korpusset af hvilke ord der ofte og typisk optræder sammen med et bestemt opslagsord. De er især nyttige for den der skal producere en tekst og slår op i ordbogen for at finde den helt rigtige formulering, men ligesom citaterne tjener de også det formål at supplere definitionerne og vise hvordan sproget bruges i praksis: Sådan udtrykker man sig typisk på dansk.

Styrken ved at have et så bredt grundlag som DDO har haft, er naturligvis at man derved undgår at fremhæve en bestemt sprogbrug (som oftest redaktørens egen), men viser sproget som det virkelig bruges. Vi kan vist

godt stå inde for at alle de kollokationer som bringes i DDO, har støtte i den faktiske sprogbrug. Til gengæld kan der være den begrænsning at nogle oplagte sprogbrugseksempler ikke er taget med fordi de ikke har været tilstrækkelig udbredte. Eksempelvis finder man *begå (et) røveri* i artiklen *røveri*, men ikke *begå (et) tyveri* under *tyveri*, og *begå et drab* står under *drab*, men *begå et mord* findes ikke i artiklen *mord*. Det kan forekomme tilfældigt, og det er det nok også. En tilsvarende søgning i et andet balanceret korpus som Korpus 2000 (www.korpus2000.dk) viser i hvert fald at *begå* optræder som typisk naboord for alle ordene *røveri*, *tyveri*, *drab* og *mord*. Man må altså leve med den slags lidt tilfældige unøjagtigheder. Det må også nævnes at vi i mange tilfælde helt har afholdt os fra at bringe sprogbrugseksempler, nemlig ved de mere sjældne ord og betydninger, fordi der har været for få forekomster til at det har været statistisk forsvarligt at bringe dem. Et eksempel på dette er verbet *gennemtrumfe*, som i ordbogens korpus "kun" har 77 forekomster (til sammenligning forekommer ordet *mord* 1327 gange). Her har redaktøren på grund af det ret lave antal absolutte forekomster undladt at bringe ellers udmærkede kollokationer som *gennemtrumfe sin vilje* eller *gennemtrumfe en beslutning* selvom de med 3-4 forekomster i korpus er forholdsvis udbredte i korpus og dermed måske også i sproget. Hvis korpus havde været to eller tre gange så stort, ville det med andre ord have ført til at vi havde kunnet bringe langt flere eksempler end det er sket.

Citater

“En ordbog uden eksempler er et skelet”. Dette fyndord tilskrives den berømte franske ordbogsforfatter Pierre Larousse (1817-75), som gik nye veje i forhold til sin samtids franske ordbøger, der i hans øjne kun var skeletter. Man kan selvfølgelig ikke klare sig uden et skelet, for det der bærer en betydningsordbog, er opslagsordene og deres definitioner, men det der giver kød på, er eksemplerne. En stor del af eksempelmaterialiet i DDO udgøres af de såkaldte “døde” eksempler der er omtalt i det foregående afsnit om kollokationer, men den mest spændende del af eksemplerne er de autentiske citater. Der er ca. 100.000 af dem i hele ordbogen, og de er netop eksempler på levende sprogbrug idet de næsten alle er taget fra ordbogens korpus med undtagelse af et par tusind, der er indsendt af opmærksomme sprogbrugere (de såkaldte spORD-hunde), hentet fra internettet eller taget fra avisdatabasen Infomedia. I begyndelsen af ordbogens redaktionsperiode var det tilladt undtagelsesvis at bruge redaktionelle eksempler, altså nogle som redaktørerne selv fandt på. Denne metode er der egentlig ikke noget galt i; den er blevet brugt i mange udmærkede ordbøger, bl.a. i Nudansk Ordbog og sågar i Ordbog over det danske Sprog, men i DDO’s redaktion besluttede vi udelukkende at bringe autentiske citater, netop for at fastholde princippet om at være korpusbaseret, at afspejle sproget som det er.

Princippet om autentiske citater betyder at der er en stor bredde i de citerede kilder – som alle angives med

årstal og fx forfatter og titel – og det fremgår af et tilfældigt opslag i bind 4, side 105, at redaktionen har været vidt omkring. Ud over traditionelle skriftsprogskilder som dag- og ugeblade (fx Politiken, Berlingske Tidende, Familie Journalen) og skøn- og faglitteratur (fx Peter Rønnov-Jessen: *En simpel historie* og en håndbog i psykologi redigeret af Mogens Hansen og Poul Erik Pagaard) er der bragt citater fra et privat brev, en blå bog fra Sønderborg Gymnasium, et magasinprogram fra DR-TV og udskrevet talesprog fra København. Tilsvarende ser man også markante forskelle i formalitetsgrad i følgende citater fra to af de ovennævnte kilder:

*den refleksive erkendelse
.. er middelbar i mod-
sætning til den intuitive
erkendelse, som er umid-
delbar håndb-psyk.90*

*han kender ingen mid-
delvej mht. alkohol-
forbrug: enten er han
ædru .. eller også er han
kanonstiv blåbogSG89*

Vi i redaktionen tror at kombinationen af autentisk citat og kildeangivelse forøger dels brugsværdien ved at oplyse om hvordan opslagsordene kan bruges “i fri dressur”, dels underholdningsværdien ved at fortælle små historier der hver især belyser den kultur og den tid de tilhører. Det er ikke så få brugere der har fortalt os at “leksikoneffekten” (“jeg kom til at læse om noget andet end det

jeg slog op”) også kan slå igennem i en trespaltet ordbog uden illustrationer!

Henrik Lorentzen (f. 1960) er senior-redaktør ved Det Danske Sprog- og Litteraturselskab.

Lars Trap-Jensen (f. 1960) er ledende redaktør ved Det Danske Sprog- og Litteraturselskab.

Den Danske Ordbog, bd. 1-6. Det Danske Sprog- og Litteraturselskab og Gyldendal, 2003-2005.

Artikler om Den Danske Ordbog i Nyt fra Sprognævnet (NfS)

Ebba Hjort: Den Danske Ordbog - en kommende kollega til Retskrivningsordbogen. NfS 1999/3, s. 1-9.

Niels Davidsen-Nielsen: En ordbog der åbner op. NfS 2004/2, s. 11-17

Omtaler af Den Danske Ordbog i Sprog i Norden (SiN) og Nyt fra Sprognævnet (NfS)

Bind 1 er omtalt i SiN 2004

Bind 2 og 3 er omtalt i SiN 2005

Bind 4, 5 og 6 er omtalt i SiN 2006

Bind 5 er desuden omtalt i NfS 2005/3, s. 14-15