

Nyhedsbrev

fra Den Danske Ordbog

Nr. 8 April 1997



Lidt om ordforråd

Som redaktør ved Den Danske Ordbog bliver man ofte mødt med spørgsmålet: "Hvor mange ord er der egentlig i dansk?". Så har man lyst til at svare: "Hvad forstår du egentlig ved et ord?". Det må man da vide før man kan beskæftige sig med hvor mange ord der er i et sprog, eller hvor mange ord en person kan, dvs. sprogets eller personens *ordforråd*.

Hvad er et ord?

Et ord er vel det der står mellem to mellemrum i en tekst (derimod kan et ord ikke være det der høres mellem to pauser i talesprog, for pauserne i et stykke talesprog passer overhovedet ikke med mellemrummene i den tilsvarende skrevne tekst). Hver gang vi mellem to mellemrum møder de samme bogstaver i den samme rækkefølge, har vi altså det samme ord. Nej, den går ikke. Det ser man let ud fra eksemplet med landmanden der blev inviteret til en fest og noget mystificerende skrev tilbage: "Hvis jeg kan så kommer jeg ikke. Hvis jeg ikke kan så kommer jeg". Det skal lige tilføjes at kommasetning just ikke var hans stærke side! To ens tegnfølger i en tekst, fx *så* adv. 'da' og *så* vb. infinitiv 'lægge plantefrø i jorden', kan betyde noget forskelligt og blive brugt forskelligt, og så må det være to forskellige ord. Sprogfolk siger da også at et ord først er et ord når det har en bestemt betydning. Og en person kan kun ordet hvis han eller hun både kan skrive/udtale ordet og kender dets betydning.

For nu at vende tilbage til tegnfølgen *så*, så kan

det også være præteritum (datid) af verbet *se*. Hvad stiller vi op med bøjningsformerne *se - ser - så - set*? Vel staves og udtales de forskelligt, og vel bruges de også forskelligt, men deres indbyrdes slægtskab gør dog at de bør regnes for fire former af det samme ord med grundformen *se* og ikke fire forskellige ord i sproget.

Hvad med sammensætninger, der jo er sammensat af to (eller flere) ord? Kender man de to ord hver for sig, kender man vel også det ord der kommer ud af det når man sætter dem sammen. Det kommer der vel ikke nødvendigvis et ord ud af som vi skal tælle med i ordforrådet. Jo, det skal vi faktisk meget tit. For en sammensætnings betydning er sjældent blot summen af de indgående deles betydning. Det første led er underordnet det sidste led i et betydningsforhold der varierer meget. En *spisestue* er fx en stue man spiser i (sted), mens en *spiseske* er en ske man spiser med (redskab). Skønt man, ud fra de indgående ords betydning hver for sig og ud fra sit kendskab til verdens indretning, i mange tilfælde kan regne ud hvad en sammensætning betyder - tag fx ordet *spisebord* -, er der ingen garanti for gevinst. Hvad betyder fx *strandfodring*? Jo, såmænd 'det at pumpe sand fra havet op på stranden for at sikre kysten mod nedbrydning'. Sammensætninger er selvstændige ord der skal tages i betragtning ved undersøgelser af ordforråd.

Disse og andre overvejelser over hvad et ord er, fører frem til at ord er det man kan slå op i en ordbog. En ordbog indeholder alfabetisk

ordnede opslagsord (herunder mange sammensætninger) i deres grundform og med hver deres betydning. Ved undersøgelser af ordforråd er det praktisk at definere ordet som opslagsordet, også kaldet det leksikalske ord.

Hvor mange ord er der i dansk?

Et fristende svar vil være: uendeligt mange, i matematisk forstand. Uendelighed er et næsten ubegribeligt begreb, i hvert fald hvis vi prøver at forestille os et uendeligt verdensrum. Bedre går det med talrækken, som vi godt ved er uendelig, oven i købet i begge ender. Uanset hvor stort et tal man tager, vil man altid kunne lægge 1 til og få et tal der er endnu større. Tilsvarende i den negative ende. Men vi har jo også ord for tallene; det er ikke bare en række cifre på papiret. Vil det så ikke sige at vi har lige så mange ord som tal, nemlig uendeligt mange? Og så er det endda kun talordene vi her snakker om. Nej, ordet for tallet 34.107.369 hører ikke til det danske ordforråd. Det er ikke et ord som sproget dansk har fikst og færdigt på lager til disposition for danske sprogbrugere. Men dansk omfatter blandt meget andet en kode der sætter dansktalende i stand til at generere ordet ud fra tallet og omvendt transformere ordet til tallet. Som sproglige byggeklodser i disse to modsat rettede processer bruger vi de talord der faktisk tilhører ordforrådet, nemlig *en, to, tre, ..., atten, nitten, tyve, tredive, fyrre, ..., hundrede, tusind, million, milliard*, osv. Det er et endeligt antal ord, og pointen er at vi ikke ud fra et endeligt antal ord kan danne et uendeligt antal ord som vi samtidig påstår tilhører det danske ordforråd.

På samme måde forholder det sig med sammensætninger. Der er et endeligt antal ord til rådighed for dannelse af sammensætninger, og kun hvis vi påstår at danske sammensætninger kan være uendeligt lange, kan vi hævde at der er uendeligt mange ord i dansk. Og det er selvfølgelig nonsens. Al respekt for H.C. Andersens *gedebukkebensoverogundergeneralkrigskommandersergent* og hans senere, mere ydmyge makker, *sporvognsskinneskiidskraberen*.

Af særlige grunde tilhører de begge to det danske ordforråd, men tilfældige eller konstruerede sammensætninger som *nyhedsbrevslæserordbogsinteresse* gør det ikke. Dét ord har aldrig før været brugt på dansk, bliver det sikkert heller aldrig mere, men det kan konstrueres og forstås af alle der behersker det danske sprogs regler for ind- og afkodning af sammensætninger.

Sprogets ordforråd er altså ikke uendeligt stort. Hvor mange ord det indeholder, bliver et spørgsmål om at afgøre hvilke ord sproget har nogenlunde færdige på lager som enheder til direkte brug af sprogbrugere uden at der skal ske nogen særlig kodning af dem. Det bliver et spørgsmål om ordenes etablerethed hos sprogbrugere. Problemet er så at det der er etableret og på lager for den ene sprogbruger, fx ordet *strandfodring* for en ekspert i kystsikring, ikke er det for en anden. Blandt andet fordi vi arbejder inden for forskellige fagområder, og fordi vi har forskellige fritidsinteresser, er vores ordforråd forskelligt i omfang og indhold, sådan som vi skal se det om lidt. Man må derfor definere det danske sprogs totale ordforråd som summen af alle de forskellige ord der findes etableret hos alle tænkelige grupper af danskere.

Det bliver et gigantisk antal, antagelig over en million ord. Det vil være umuligt at lave en sådan fuldstændig komplet ordbog over dansk. Det vil heller ikke være interessant fordi mange af ordene er så specielle at de kun kendes af et meget lille antal danskere. Eksisterende og kommende ordbøger må, afhængigt af ordbogens størrelse, begrænse sig til de mere eller mindre centrale dele af ordforrådet: almenordene og de mest udbredte fagord, eller for fagordbøgernes vedkommende kun de mest udbredte fagord. Her er nogle tal der angiver hvor mange opslagsord der findes i nogle kendte ordbøger (ifølge ordbøgernes egne opgivelser):

Ordbog	Antal opslagsord
Oxford English Dictionary (1992)	over 500.000
Webster's Third New International (1961)	over 450.000
Blinkenberg & Høybye: Dansk-Fransk Ordbog (1991)	mere end 172.000
Retskrivningsordbogen (Dansk Sprognævn, 1996)	ca. 64.000
Retskrivningsordbog (Gyldendal, 1996)	ca. 80.000
Nudansk Ordbog (1996)	over 100.000 ord, konstruktioner og faste vendinger
Den Danske Ordbog	ca. 125.000 ord, ordforbindelser og sammensætninger

Hvor mange ord kan du?

Det kan vi naturligvis ikke svare på, men hvilken størrelsesorden tror du det er? Prøv at gætte! Her vil vi fortælle lidt om hvordan man kan måle en persons ordforråd, og hvilke resultater målingerne giver. Fra en meget stor ordbog, fx Blinkenberg & Høybyes Dansk-Fransk Ordbog, vælger man 100 statistisk helt tilfældige ord, og så tester man om forsøgspersonen kan ordene, ét for ét. At kunne et ord vil sige at man omgående genkender ordet og på én eller anden måde kan give udtryk for at man i forvejen kender den rigtige betydning (eller én af betydningerne). Hvis forsøgspersonen fx kan 25 ord af de 100 ord, dvs. en fjerdedel, ganger man op og siger at vedkommende har et ordforråd på en fjerdedel af ordbogens 172.000 ord, dvs. 43.000 ord.

Det er antagelig et større tal end det du gættede på før, men faktisk viser undersøgelser både i Danmark og i udlandet at mange voksne har et ordforråd på 50.000-70.000 ord, og at et barn i 1. klasse allerede har et ordforråd på 5.000-10.000 ord. Det er klart at der er en sammenhæng mellem alder og ordforråd, ikke blot sådan at ordforrådet vokser voldsomt i løbet af barndommen, men også i kraft af at det hos den voksne stiger med alderen, dog naturligvis langsommere. Der kan desuden påvises sammenhænge mellem på den ene side ordforråd og på den anden side viden, intelligens og læsning. Interesserede læsere henvises i øvrigt til Lars Brinks spændende artikel "Ordforrådets omfang og vækst" i Mogens Jansen og Jørn Lund

(udg.): "Børnenes sprog - sprogene omkring børn" (1979). Artiklen punkterer eftertrykkeligt en skrøne om at en pramdrager på Themsen kan 300 ord, og hvis man taber noget tungt på hans fødder, får man dem alle sammen at høre!

Af Kjeld Kristensen

Nyhedsbrevet er redigeret af Kjeld Kristensen.
 Adresse: Den Danske Ordbog, Københavns Universitet, Njalsgade 80, 2300 København S.
 Tlf.: 35 32 90 00.

Mennesker, maskiner og ord ...

En ordbog bedømmes ikke alene på kvaliteten af de artikler man kan slå op i ordbogen. Den bedømmes også på *hvilke* ord der kan slås op, og måske navnlig hvilke ord der *ikke* kan slås op i ordbogen. Hvis man gentagne gange oplever at måtte søge forgæves efter et ord, bliver man forståeligt nok irriteret. Det er kun rimeligt at kræve af en ordbog at man kan finde de mest almindelige ord i den.

amatøragtig, babytøj, brudevals, bykort, dersens, havegang, indvirken, kødsovs, langlemmet, litervis, mavemuskel, møgunge, nettohusleje, parmesan, piberyger, pikhoved, racerbane, rygvind, selvdisciplin, syvtommersøm, velsmurt er eksempler på almindelige ord som de fleste danskere kender og indimellem også bruger. Det er vel ikke specielt nye ord eller betegnelser for nye fænomener af den type som vi bl.a. modtager belæg på fra vores spORDhunde. Mange af dem er sammensætninger, men deres betydning kan langtfra altid gennemskues selv om man kender betydningen af de ord de er sammensat af. Det er altså ord som mange vil forvente at kunne slå op og finde beskrevet i en ordbog. Men det er, pudsigt nok, netop eksempler på ord som man faktisk hverken kan slå op i Retskrivningsordbogen, i Ordbog over det danske Sprog, i Nudansk Ordbog, i Vinterberg og Bodelsens store dansk-engelske ordbog eller i Blinkenberg og Høybyes store dansk-franske ordbog. Det er eksempler på ord som, måske ved tilfældighedernes puds, er blevet oversat ved tilblivelsen af disse ordbøger, måske især ved tilblivelsen af Ordbog over det danske Sprog, hvis ordudvalg ligger til grund for mange senere ordbøger.

Det er ikke fordi der er tale om dårlige ordbøger. Tværtimod repræsenterer de noget af det ypperste inden for dansk ordbogstradition. Men man må huske på at i tiden før edb-alderen havde redaktørerne nogle ganske andre arbejdsbetingelser end vi har i dag. Dengang skete udvælgelsen af ordstoffet manuelt. Redaktionens hjerte var et stort arkiv med kartotekskasser fyldt med belæg på ord som redaktørerne havde

iagttaget og registreret på små sedler. Det siger sig selv at næsten lige meget hvor dygtige og omhyggelige redaktørerne har været til at registrere deres sproglige iagttagelser, er det uundgåeligt at et enkelt ord er smuttet nu og da. Ord som blev oversat ved en tilfældighed.

I dag er vi i den heldige situation at vi har computere til at hjælpe os med at sikre at vi ikke overser almindelige ord. I vores tekstsamling har vi et meget stort antal tekster fra vidt forskellige genrer i elektronisk form. Ved hjælp af et særligt edb-program som er knyttet til den elektroniske tekstsamling, gør computeren os automatisk opmærksom på ord som forekommer med en vis hyppighed i vores tekster, sådan at vi tvinges til at vurdere om det er ord der bør optages i Den Danske Ordbog. Det kan godt være at vi vælger at se bort fra ordet, men det smutter ikke ved en tilfældighed. Det er endnu et af de områder hvor moderne ordbogsarbejde drager nytte af edb-teknologien.

Her er nogle flere eksempler på ord som heller ikke kan slås op i de nævnte værker, men som sikkert kommer til at stå i Den Danske Ordbog:

bilmotor, bilos, bryllupsklokker, cigaretrøg, dagpengesystem, danskundervisning, d'herrer, drømmeprins, efterkrigsår, eftervarme, elvarme, engelsklærer, festtøj, fiberskade, finansverden, firehjulstrukket, fødselsdagssang, førstesal, gammelmandssyn, glatnakke, græskortodoks, halvskaldet, herr, historiebog, højdramatisk, kirsebærtræ, klasselokale, knytænveslag, kræftformer, køkkengulv, mandehørm, markedsøkonomisk, målchance, opslagsbog, passagersæde, pladestudie, privatøkonomi, renteudgift, reservemålmand, skiferie, slutfacit, stamværtshus, standarddans, stillingsopslag, telefonregning, testikelkræft, titusindvis, trommesolo, underlivsundersøgelse, urutineret, valgaften, veludhvilet, vittighedstegning, værtshusslagsmål ...

Af Lars Trap-Jensen