

# Danmarks Sprog-museet

Sprog-museet.dk

Redaktør: Ole Stig Andersen

## Den Danske Ordbog på nettet

af Lars Trap-Jensen, den 21. april 2010

Siden november 2009 har Den Danske Ordbog været tilgængelig på nettet.

Redaktionen af Sprog-museet

har i den anledning bedt mig give en omtale af ordbogen, dens omfang og anlæg generelt og mere specifikt beskrive hvordan den fungerer i netversionen.



Den Danske Ordbog indskrives sig i traditionen af danske modersmålsordbøger, dvs. ordbøger der registrerer og forklarer danske ord ved hjælp af andre danske ord. Den primære brugergruppe er derfor naturligt nok personer med dansk som modersmål, men også udlændinge der lærer dansk på højere niveau, vil have glæde af ordbogen. Ordbogen beskriver det almindelige ordforråd i moderne dansk fra ca. 1955 til i dag, med hovedvægten på en kerneperiode 1983-1992. Den giver besked om ordenes **stavning, bøjning, udtale, oprindelse, betydning og brug**. Og den viser hvordan ordene forbindes med andre ord i grammatiske konstruktioner, faste udtryk og orddannelser.

### Baggrund og omfang

Den Danske Ordbog udkom som papirordbog i seks bind 2003-2005. I alt registrerer den trykte ordbog 100.000 ord, hvoraf de 62.000 behandles som hovedopslagsord med egen artikel, mens de resterende ca. 38.000 ord indgår i hovedartiklerne som eksempelmateriale til at vise almindelige afledninger, sammensætninger, synonymer, antonymer m.v.

Papir og skærm er forskellige publiceringsmedier, og netudgaven adskiller sig derfor på væsentlige punkter fra den trykte ordbog, også indholdsmæssigt. En vigtig forskel er at pladsøkonomi spiller en vigtig rolle når man redigerer trykte ordbøger, men næsten ingen ved publicering på nettet. Omfangsmæssigt betyder det at netudgaven indeholder både flere og færre opslagsord end papirudgaven.

Der er færre opslagsord i netudgaven fordi en masse artikler har kunnet nedlægges, nemlig alle dem der i den trykte ordbog udelukkende tjener til at oplyse om at ordet skal findes et andet sted end dér hvor man har slået op. Ellers kunne brugeren måske tro at det pågældende ord slet ikke fandtes i ordbogen. Hvis man eksempelvis slår op i den trykte udgave under *acryl*, får man at vide at denne stavemåde ikke er officielt godkendt, og henvises derfor til den officielle form *akryl*. I netudgaven kan man godt slå de to led sammen til ét så brugeren ledes til det rette opslagsord straks efter indtastningen. Ved simple slåfejl eller almindelige stavfejl får man i netudgaven hjælp af funktionen "Mente du: ..." lige under søgefeltet, der kommer med alternative forslag

til opslagsord der minder om det indtastede.

Til gengæld er der flere opslagsord med fordi alle de mange eksempelord som i den trykte ordbog kun optræder i artikler om andre opslagsord, har fået deres egen artikel i netudgaven. På den måde kan man hurtigt finde ud af om et ord er taget med i ordbogen, og som nævnt betyder den ekstra plads ikke længere noget. De fleste af den slags artikler har indtil videre kun oplysning om opslagsform, ordklasse og bøjning, men mangler betydningsdefinitioner, citater og de andre oplysninger der normalt optræder i en ordbogsartikel.

Endelig er der redigeret helt nye artikler siden den trykte ordbog udkom, og der kommer hele tiden flere til. Modsat den trykte ordbog ændrer netudgaven sig altså løbende, og præcise oplysninger om omfanget bliver derfor hurtigt forældet. I marts 2010 var der ca. 95.000 opslagsord i ordbogen.

## En korpusbaseret ordbog

Den Danske Ordbog er en *korpusbaseret* ordbog, faktisk den første danske almenordbog der bygger på et elektronisk tekstkorpus. Den Danske Ordbogs korpus udgøres af en stor samling digitale tekster indsamlet i perioden 1983-1992 og sammensat så de så vidt muligt afspejler dansk sprog i dets mange afskygninger. Der indgår aviser og romaner, fagbøger og tidsskrifter, tekster produceret både af professionelle og amatører, på skrift såvel som i tale. Der er med vilje ikke gjort forsøg på at indfange det særlig "korrekte", "smukke" eller på andre måder eksemplariske sprog. Meningen er at registrere sproget netop som det tales og skrives af et bredt udsnit af danske sprogbrugere. Ordbogens korpus indeholder derfor også en del andre tekster end dem man almindeligvis støder på fx ved søgninger på internettet, ligesom forholdet imellem dem er et andet.

Det sprog der på den måde fremkommer, afspejler på godt og ondt moderne dansk i slutningen af 1900-tallet. Her findes journalisters og andre fagfolks professionelle tekster og ordkunstneres smukke og poetiske sprog side om side med skoleelevers og folkepensionisters sprog, talesprog fra radio og tv-udsendelser, foruden hverdagens anonyme og upåfaldende tekster fra pjecer og tilbudsaviser, ja endda højtalerannonceringer fra S-togs-stationer er der blevet plads til. Endelig er der lagt vægt på også at få fat i helt almindelige menneskers tekstproduktion fra private breve og dagbogsoptegnelser. Det er sproget i alle disse tekster der er analyseret og beskrevet i ordbogen; i alt indeholder ordbogens korpus næsten 40 mio. løbende ord.

Den Danske Ordbogs citater har plads til både højt og lavt sprog. Her er et par eksempler: artiklerne *smøge* og *spor*.

### 1.a OVERFØRT komme af med noget tyngende eller besværende

SYNONYM smytte

GRAMMATIK NOGEN smøger NOGET +ADVERBIAL

det er som om de først vågner nu og smøger alle fortrædelighederne af sig og retter på håret KIRifb92

Fra artiklen *smøge*

## 2.a nummereret jernbanespor på en station hvor tog ankommer og afgår

SE OGSÅ [perron](#)

Tog til Høje Tåstrup afgår fra spor 3. Bemærk sporændring, spor 3 [S-togs-ann91](#)

Fra artiklen *spor*

Ordnes hyppighed i korpus har haft betydning for udvælgelsen af opslagsord til ordbogen, og beskrivelsen af deres brug er blevet til ved at analysere de eksempler der findes i korпустeksterne. De betydninger der er taget med, har redaktørerne fundet eksempler på i korpusset, ligesom de citater der er udvalgt til illustration af betydningerne, i alt næsten 100.000, alle er hentet fra autentisk sprogbrug. Også oplysninger om en bestemt sprogbrug (*især yngre eller ældre, især Jylland, sjældent osv.*), grammatiske mønstre og ordforbindelser stammer fra undersøgelser af sproget i de mange korпустekster.

Fordi der er så tæt en sammenhæng mellem ordbog og korpus, vil vi i netudgaven gerne knytte dem nærmere sammen. En del af Den Danske Ordbogs korpus indgår i KorpusDK, som findes på samme sted som ordbogen, nemlig som en del af <http://ordnet.dk>, og mellem disse to resurser er der skabt forskellige forbindelser.



### Relaterede søgninger

#### Teksteksempler

[spor](#)

[spor, sb.](#)

#### Naboord

[spor](#)

[spor, sb.](#)

#### Ordbog over det danske Sprog

[spor](#)

[spor, sb.](#)

Tydeligst ser man det i venstrespalten, hvor der vises eksempler på forskellige “Relaterede søgninger”. Når man befinder sig i en ordbogsartikel, kan man klikke på *Teksteksempler* og se forekomster af opslagsordet i korpus.

Også inde i artiklen om opslagsordet kan man slå op i korpus: De korte eksempler på typisk sprogbrug der gives før citaterne, er hentet fra korpusundersøgelser, og hvis man gerne vil se eksempler på sprogbrugen, skal man klikke på det lille “k”-ikon der står efter eksemplet. Så foretages der en søgning efter den pågældende

ordforbindelse (eller *kollokation* med en faglig betegnelse) i korpus. I artiklen *spor* finder man på den måde: *på rette spor, afgørende spor, slette alle spor, dybe spor, synlige spor, sætte sine spor, skifte spor, tilbage på sporet.*

interessegruppe hævder, at reformerne er kørt af **sporet**, og hvis der ikke finder en alvorlig korrektion sted, vil kulturen. Hans indsats på dette område har sat sig varige **spor**- bl.a. i form af de kulturelle sommerarrangementer i København. Siden vibrerende, sniffende tå. Bill Holmberg trækker slimede **spor** efter sig i rollen som den sortrandede, djævelske skibslæge fra sexliv føler de sig mere lykkelige. Cesira er ikke **spor** intellektuel. Som Adriana har hun ej heller læst lødige litteratur. desto mindre en privat efterforskning. De kommer også på **sporet** af forbryderne og er lige ved at indkredse dem, da enkefru Løwenskjolds forsvundne ægtemand, som han har **sporet** til Italien. SÅ den gamle ræv planlægger, for enkefruens penge, at han på nuværende tidspunkt ikke har besluttet **spor** at gå efter, har han noget at have optimismen i. brande. Langsomt og sikkert fik Per Casper og co. sig **sporet** ind på pyromanen, og fire måneder efter den første af leder, advarer mod blot at lade intifadaen fortsætte i samme **spor**. Adwa er en stor provinsby i Tigre. Den ligger midt kræver diskretion- konkurrenterne må ikke ledes på **sporet** af muligheder, der findes. Nakskov-virksomheden er blandt de ypperste rent P. Jensen elev i slutningen af tyverne. Det vakte ikke **spor** forundring, at Carl P. Jensen, som var formand for sin hun ild og den absurde indbyrdes " kamp " efterlader kun et **spor** af lunken hadefuldhed. Her har et net af vaner holdt

## Konkordans fra korpusDK

Klikker man på "Naboord" i venstrespalten, får man vist en samlet oversigt over de ord der typisk står ved siden af opslagsordet i korpusteksterne, henholdsvis til venstre og højre for ordet.

### Venstre kontekst

▲	ordform	ordklasse	score	signifikans
1	sletter	v	11.10	██████
2	slette	v	10.86	██████
3	tydelige	adj	10.04	██████
4	efterlade	v	9.09	██████
5	varige	adj	8.86	██████
6	dybe	adj	8.70	██████
7	synlige	adj	8.59	██████
8	krydser	v	8.57	██████
9	rette	adj	8.28	██████
10	efterladt	v	8.26	██████
11	efterlader	v	8.10	██████
12	blodige	adj	7.97	██████
13	intet	det	7.58	██████
14	ethvert	det	7.54	██████
15	sæt	v	7.10	██████
16	skiftede	v	6.87	██████

### Højre kontekst

▲	ordform	ordklasse	score	signifikans
1	skræmmer	v	10.62	██████
2	forbavset	v	8.67	██████
3	sneen	n	7.77	██████
4	sjovt	adj	7.57	██████
5	jagten	n	7.49	██████
6	ked	adj	7.41	██████
7	nervøs	adj	6.51	██████
8	rart	adj	6.46	██████
9	imod	prp	5.81	██████
10	lyst	n	5.72	██████
11	sagen	n	5.25	██████
12	bange	adj	5.13	██████
13	anderledes	adj	4.81	██████
14	historien	n	4.69	██████
15	efter	prp	4.58	██████
16	glad	adj	4.57	██████

## Liste over naboord til *spor* efter hyppighed

Et opslag i ordbogen kan på den måde lede videre til en søgning i korpus efter flere eksempler med et bestemt ord eller udtryk. Omvendt kan den der befinder sig på en korpusside, slå ordet op i ordbogen, og sådan kan man bevæge sig frem og tilbage mellem de to resurser.

I det følgende ser vi på hvordan DDO på nettet fungerer, med vægt på de dele der adskiller sig fra det trykte værk.

## Søgning

De fleste søgninger der foretages i DDO, er simple søgninger hvor der indtastes et

bestemt opslagsord og trykkes “Retur”. Hvis der kun er ét opslagsord der matcher det indtastede, vises den tilhørende ordbogsartikel i sidens midterfelt, og skulle der være flere opslagsord der matcher, vises den første af disse i midterfeltet, mens de øvrige muligheder vises i fanebladet “Søgeresultat” til højre på skærmen.

Som noget nyt kan man derudover – også i højre side af skærmen, under søgeresultatet – se en liste over de flerordsforbindelser, faste udtryk, det indtastede ord indgår i. For overblikkets skyld er listen som udgangspunkt lukket, men den kan nemt åbnes med et klik på det lille plustegn, og man har derfra mulighed for at klikke sig direkte frem til et udtryk frem for at skulle lede gennem en længere artikel. Man behøver heller ikke længere spekulere over om et udtryk som *gå gennem marv og ben* skal findes under **gå**, **marv** eller **ben**. I

netudgaven findes udtrykket både selvstændigt, hvis man indtaster hele udtrykket – eller bare nogle af ordene – i søgefeltet, og under hvert af de indgående ord i udtrykket. Det betyder heller ikke noget om man indtaster en bøjet form af ordet, det bliver fundet alligevel. Det er gjort så brugeren også får et match ved søgning på de mange udtryk der kan bøjes (kulden *gik* gennem marv og ben, de var *store* i slaget) eller varierer i form (komme i *tanke/tanker* om, skære *ansigt/ansigter*, i *fuld/fuldt* alvor).

I det hele taget er der gjort en del ud af at lette søgning og navigering på siden. Det nævnte plustegn hhv. minustegn kan bruges til at åbne og lukke for en artikels hovedafsnit. Hvis man kun er interesseret i én bestemt type oplysning, giver det bedre overblik hvis de øvrige oplysningstyper er skjulte. Hvis der er tale om lange artikler med mange betydninger og faste udtryk, kan det være en idé at finde den ønskede betydning ved hjælp af indholdsfortegnelsen, der foldes ud med et klik på bjælken “Vis overblik” øverst i artiklen. På samme måde kan man få et hurtigt overblik over alle de faste udtryk der er registreret for et opslagsord.

Når man skal læse en artikel, kan man vælge mellem to forskellige præsentationer af den: “Lang visning” og “Kort visning”. Man vælger sin foretrukne visning ved skift mellem de to faneblade øverst i artiklen. I den lange visning ser man alle de oplysninger der er registreret for opslagsordet, mens den korte visning skaber et bedre overblik ved at skjule nogle af de oplysninger som mange godt kan undvære. I den korte visning ser man derfor ikke oplysninger om udtale, etymologi, beslægtede ord, grammatiske oplysninger, korte eksempler (kollokationer), og der vises højst ét citat for hver betydning.

Også andre steder i artiklen er der gjort forsøg på at lette forståelsen af den traditionelle ordbogs komprimerede stil. De fleste forkortelser er således i netudgaven skrevet ud, og visse svære ord er forsynet med en stiplet understregning. Den viser at man kan få en boks med yderligere forklaring frem ved at holde cursoren hen over ordet, en såkaldt mouse-over. På samme måde kan man få udfoldet hvad kilden til et citat er. I den trykte ordbog er kildeangivelser altid forkortede og noget kodeprægede af hensyn til

The screenshot shows a search interface with a search bar containing the word "spor" and a "Søg" button. Below the search bar, there are navigation links: "Mente du: spore | sport | spir ...vis flere". There are two tabs: "Søgeresultat" (selected) and "Alfabetisk liste". Under "Søgeresultat", there is a section "Opslagsord (1)" with a minus sign, containing the result "spor sb.". Below that is a section "Faste udtryk (8)" with a minus sign. It contains an information icon and the text "faste udtryk der matcher søgningen". Below this are several items with right-pointing triangles: "blindt spor → blindspor", "følge i nogets spor", "have (holde el. sætte el. ...) snuden el. næsen i sporet", "ikke spor", "krydse nogens spor", "køre (løbe el. ryge el. ...) af sporet", "på sporet", and "sporene skræmmer".

pladsøkonomien, men i netudgaven kan man få vist den fuldt udfoldede kilde som mouse-over. Man behøver altså ikke længere spekulere over hvad *LokAFredHumKar192* mon betyder, for i det lille vindue står “Lokalavisen for Fredensborg-Humlebæk og Karlebo kommuner, 1992”.

Det samme virkemiddel er taget i brug i orddannelses-delen, hvor vi har forenklet listen med sammensætninger. I den trykte ordbog er sammensætningerne ordnet i grupper efter den betydning de lægger sig til, men i netudgaven har vi for overblikkets skyld kun én alfabetisk ordnet liste med sammensætninger. Til gengæld kan man med mouse-over-funktionen få vist hvilken betydning en bestemt sammensætning eventuelt er dannet til.

## Beslægtede ord

En ny oplysningstype i netordbogen er “Beslægtede ord”. Her kan man se hvilke andre ord der minder om opslagsordet i forskellige henseender. Oplysningerne i “Beslægtede ord” stammer fra projektet *DanNet*, der går ud på at opbygge et hierarkisk system hvor begreberne er ordnet efter over- og underbegreb, idet der samtidig gives andre semantiske oplysninger om fx genstandes dele, deres formål og typiske anvendelse, personers beskæftigelse m.m. Oplysningerne i *DanNet* er blevet til på grundlag af Den Danske Ordbogs opslagsord og betydningsbeskrivelser, og derfor er det oplagt at de nu – i videreforarbejdet form – vender tilbage til ordbogen med ekstra oplysninger om ordenes betydningsmæssige relationer til andre ord.

Der opereres med tre forskellige betydningsrelationer: overbegreber (mere generelt), underbegreber (mere specifikt) og søsterbegreber (ord med samme overbegreb). Eksempel: Slår man ordet *kringle* op, får man under “Beslægtede ord” at vide at overbegrebet er *kage*, og at *kringle* har underbegreber som *galopkringle* og *fødselsdagskringle*. Blandt søstertermene, altså andre kager, finder man et større bagersortiment med bl.a. *berlinerpfannkuchen*, *julekage*, *smørkage* og *brioche*. *Søsterkagen* er ikke lige til at finde, men den viser sig at være datter af *formkagen* sammen med fx *gulerodskage* og *sandkage*. Som det fremgår, kan “kagen skæres” på forskellig måde, og går vi fx til naturvidenskabelige og tekniske områder, er det sandsynligt at fagfolkene ikke vil inddele verden på samme måde som leksikograferne. Men det er netop en pointe ved den måde *DanNet* er skruet sammen på, at inddelingen så langt det er muligt, følger almensproget og ikke fagsproget.

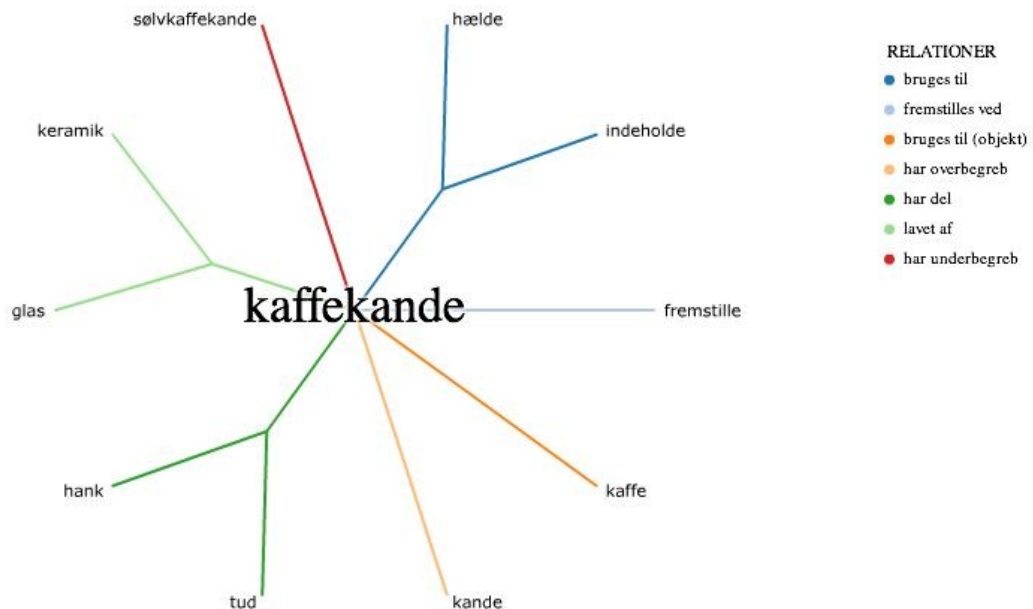
Indtil videre er “Beslægtede ord” stadig lidt af et eksperiment, og den er derfor blevet forsynet med et lille mærke “beta” mens der fortsat eksperimenteres med visningen af de forskellige relationer. Endelig er der et link til en ekstern resurse, <http://andreord.dk>, hvor man kan se en grafisk præsentation af den pågældende betydning og dets relationer til andre ord.

I eksemplet fra [andreord.dk](http://andreord.dk) kan man se at *kaffekande* har overbegrebet *kande*, mens det har *sølvkaffekande* som underbegreb. Man kan se at det bruges til *kaffe*, kan være lavet af *glas* eller *keramik*, kan have en *tud* og en *hank* og bruges til at *hælde* og *indeholde* noget i. Hvis man vil se andre ord med *kande* som overbegreb, altså sidebegreber som *termokande*, *mælkekande* og *tinkande*, må man først klikke på overbegrebet *kande*, så det positionerer sig i midten. Søsterbegreberne vil da være alle de underbegreber der er til *kande*. I nogle tilfælde findes der rigtig mange beslægtede ord, navnlig underbegreber og søsterbegreber, og i det tilfælde udregnes der derfor en score for lighed som bruges til at rangordne ordene med.

## kaffekande

SUBSTANTIV

kande brugt til (servering af) kaffe



Visuel fremstilling af 'kaffekande's relationer til andre ord

### Beskrivelse og normering

Som en korpusbaseret ordbog indgår Den Danske Ordbog i traditionen for deskriptive ordbøger. Det betyder at opslagsordene er udvalgt og beskrevet på grundlag af det materiale som viser sig i korpus. Men deraf følger ikke at alt hvad der står i ordbogen, er den objektive sandhed om sproget. Til ordbogsarbejdet hører også at foretage skøn og træffe beslutninger, ja faktisk giver det ikke mening at betegne en ordbog som rent deskriptiv, for der er adskillige oplysninger i en ordbog som falder uden for deskriptionen. Antal opslagsord, opdelingen i betydninger og underbetydninger, oplysninger om synonymer og antonymer eller sprogbrugsoplysninger som *nedsættende*, *gammeldags*, *uformelt* eller *dialektpræget* er ikke beskrivelser af iagttagelige og objektive fænomener. Kommentarerne er ofte mere vurderende end rent deskriptive; de kan være mere eller mindre velanbragte, karakterisere en tendens mere eller mindre overbevisende, men de er næsten aldrig udelukkende sande eller falske. Når man laver ordbøger, kan man altså ikke altid være rent beskrivende. Leksikografen er vurderende, kommer med – direkte eller indirekte – advarsler og anbefalinger og er normerende når det analyserede materiale skal præsenteres med tanke på brugeren. Ordbogsarbejdet rummer derfor både deskriptive og normative aspekter.

De opslagsord der er taget med i Den Danske Ordbog, er opført i den form de har i Retskrivningsordbogen fra Dansk Sprognævn (3. udgave, 2001). Man kan altså godt bruge Den Danske Ordbog som retskrivningsordbog. Dog skal man være opmærksom på nogle særlige forhold:

Der findes ord i Den Danske Ordbog som ikke står i Retskrivningsordbogen. Her har redaktionen selv fastlagt en opslagsform, som regel den form der er mest udbredt i sproget. Hvis det har været muligt, træffes valget på grundlag af de ord som Dansk Sprognævn har taget stilling til: Når Retskrivningsordbogen har formerne *linje*, *linjefag*, *linjeføring* osv., står der i Den Danske Ordbog *linjeafstand*, *linjedommer*, *linjeforløb* osv. selvom de sidstnævnte ord ikke er taget med i Retskrivningsordbogen. I andre tilfælde hvor der kan være større tvivl, har redaktionen rådført sig med Dansk Sprognævn inden opslagsformen blev fastlagt.

Der findes stavemåder som er etablerede inden for et bestemt fagligt område, men som er opført i Retskrivningsordbogen i andre former. Her vælger Den Danske Ordbog at opføre ordene sådan som de står i Retskrivningsordbogen, men anfører også den fagligt etablerede, men altså uofficielle form. Uofficielle former står med en svagere grå farve i en rød stiplede kasse og har en kommentar der forklarer formen, som fx artiklen **ætylen**: *uofficiel stavemåde, almindelig blandt kemikere: ethylen og ethen*.

Der findes variantformer der ikke står i Retskrivningsordbogen, men som alligevel må regnes for acceptable dannelser ifølge de sproglige regler for dansk. Fx er formen *halmpresse* næsten lige så udbredt i sproget som *halmpresser* selvom kun den sidstnævnte form er taget med i Retskrivningsordbogen. *Halmpresse* er sammensat helt i overensstemmelse med reglerne for orddannelse (Retskrivningsordbogen har fx optaget dobbeltformen *ballepresse* eller *ballepresser*), og derfor har Den Danske Ordbog valgt at tage denne form med som en variant af *halmpresser*. Men i modsætning til de officielle dobbeltformer, som blot anføres med et *eller* imellem, kan man se at variantformen ikke er en officiel dobbeltform: Varianten står med mindre skrift og er forsynet med en kommentar: *også i formen: halmpresse*.

## Brugerrespons

Som så mange andre hjemmesider er Den Danske Ordbog interesseret i feedback fra sine brugere. Vi vil gerne forbedre sitet dér hvor der er brug for det, og har naturligvis selv nogle idéer til hvordan det kan gøres. Men det er vigtigt for os at vide om brugerne nu også finder det de har brug for, og hvordan de i det hele taget anvender ordbogen. Analyser af logfiler kan hjælpe ved at vise søgemønstre og ord der ofte søges efter, og det kan vi fx bruge når vi skal udvælge nye opslagsord. En liste med de hyppigst søgte opslagsord findes på ordbogens forsider. Desuden kan man som bruger selv indsende forslag til nye opslagsord ved hjælp af funktionen "Send et ord" i venstre spalte. Forslagene registreres i en database som Det Danske Sprog- og Litteraturselskab har oprettet og driver i fællesskab med Dansk Sprognævn. Nogle af de indsendte forslag til nye ord kan ses på ordbogens forsider. Og ellers er der mulighed for at kontakte redaktionen med ris og ros, fejl, kommentarer og forslag på redaktionens e-adresser: [ddo@dsl.dk](mailto:ddo@dsl.dk) eller [ordnet@dsl.dk](mailto:ordnet@dsl.dk). Det vil jeg afslutningsvis også opfordre denne artikels læsere til at benytte sig af.

**Lars Trap-Jensen**, ledende redaktør  
Det Danske Sprog- og Litteraturselskab