

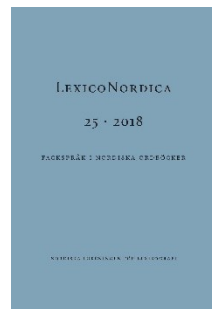
LexicoNordica

Titel: *Den Danske Ordbog på nettet - en undersøgelse af version 3.0*

Forfatter: Ken Farø, Lene Vinther Jensen

Kilde: LexicoNordica 25, 2018, s. 215-232

URL: <https://tidsskrift.dk/index.php/lexn/issue/archive>



© 2018 LexicoNordica og forfatterne

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.

Den Danske Ordbog på nettet – en undersøgelse af version 3.0

Ken Farø & Lene Vinther Jensen

This review article takes a selective look at the electronic corpus-based Danish dictionary *Den Danske Ordbog*. We aim at a descriptive and critical – yet fair – analysis of this major pioneering work within Danish monolingual lexicography. Focusing on the online version and some of its new features, we try to find out whether this lexical tool meets the needs in terms of coverage of both recent and more general lexis. We also suggest a new way to perceive corpus-based dictionaries and the possible consequences of their inclusion policies.

1. Indledning

Den Danske Ordbog (DDO) er et *flagskib*. Både for Det Danske Sprog- og Litteraturselskab (DSL) og for den danske ordbogsverden som helhed. Ifølge ordbogen selv er et flagskib (bl.a.) betegnelsen for “et projekt eller et produkt, der er det største eller mest betydningsfulde for fx et firma”. Selvfølgelig er der forgænger, *Ordbog over det Danske Sprog* (ODS, 1918-55)¹, som DDO med hensyn til omfang (6 vs. 33 bind) ikke når til sokkeholderne. Men DDO vil og kan noget andet end sin forgænger: Ud over at være stærkt dynamisk og interaktiv, er den videnskabeligt radikal. I hvert fald på de fleste punkter. Den er med afsæt i KorpusDK i høj grad *korpusbaseret*, ikke bare *korpusstøttet* som en del andre søsterordbøger (fx *Duden*-ordbøgerne og *Nudansk Ordbog*). På dette punkt ligner den fx *Macmillan English Dictionary* (MED). Vi vil imidlertid pege på et par punkter, hvor den ikke er radikal

1 ODS dækker det danske ordforråd fra ca. 1700 til 1950.

nok. Desuden vil vi forelægge en kættersk tanke til udviklingen af fremtidens korpusbaserede ordbog, nemlig en anderledes tilgang til frekvensprincippet for lemmaoptaget end hidtil. Vores hovedpunkter er: 1) nye features, 2) dobbeltformer og beslægtede variansfænomener, 3) det semantisk/pragmatiske felt ‘tak’², 4) ordbogens lemmabestand, og 5) muligheden for justering af algoritmen for inklusion.

Eftersom især den trykte version har været genstand for utallige omtaler siden 2003 (jf. nedenfor og litteraturlisten), vil vi fokusere på onlineudgavens funktioner og indhold. Titlen på denne artikel skal forstås sådan, at vi opfatter bogudgaven (bind 1-6, udgivet 2003-05) som version 1.0, den første selvstændige onlineversion (2009) som 2.0, og den foreliggende som 3.0. (2018). Artiklen er notabene skrevet inden opdateringen af DDO i juni 2018.

2. Tidligere undersøgelser og omtaler

I 2003-2018 er der udkommet en stor mængde omtaler af DDO. Nogle af de mest produktive – og samtidig mest kritiske – afsendere har været Århus-skolen med afsæt i sin “funktionslære”, både fordi skolen har udgjort et stærkt fagligt og ambitiøst fællesskab, men også fordi den nok i nogen grad havde et horn i siden på det konkurrerende projekt, der havde modtaget et tocifret millionbeløb i stats- og fondsstøtte, hvormed den var en udfordring for *DanskOrdbogen*, som var Århus-skolens projekt. Nogle af de punkter, der er blevet kritiseret, er et angiveligt for lille og oppustet lemmaomfang (Bergenholtz & Vrang 2004) og kildehenvisningernes funktion (Leroyer 2005) – se også Pálfi (2005) for en tidlig (og ikke helt upartisk) sammenfatning af omtalerne.

2 Vi arbejder med dette fænomen i dansk-tysk perspektiv. Jf. også Faro (2018).

3. Nyt i Den Danske Ordbog online

3.1. “Ord i nærheden”

Tilgangen “Ord i nærheden” er baseret på DDO’s søsterordbog *Den Danske Begrebsordbog* (DDB 2014), hvis opslag igen alle er forbundne til ord, der er optaget i DDO (DSL 2018). *Ord i nærheden* er en interessant feature, hvis nytteværdi nok ikke falder alle brugere ind lige med det samme. At den både kan dække over synonymer, over- og underbegreber og endda antonymer, var vi nok ikke selv kommet på. Når funktionen endvidere dækker over frasermer (fx henvisningen fra *notabene* til *kort sagt*), begynder betegnelsen “ord i nærheden” at blive misvisende. Den rumligt konnoterende term kunne måske erstattes af en, der i højere grad signalerer, hvad der er tale om, nemlig semantiske strukturer af ret forskellig art. Det kunne fx være “ord og udtryk med relaterede betydninger”. Det lyder måske mindre friskt, men man kan ikke få i både pose og sæk.

3.2. Leksikalske opdateringer

Siden sin onlinelancering er ordbogen med jævne mellemrum blevet tilføjet nye ord og udtryk (jf. afsnit 6). Det foregår typisk bunkevist frem for løbende, og altid under stor mediebevågenhed. Om den rykvise påfyldning skyldes tekniske betingelser, eller om det er af reklamehensyn, ved vi ikke. I hvert fald får processen på den måde karakter af en begivenhed, og medierne beretter villigt om, at dette eller hine ord nu er blevet optaget i dansk, som altså sættes lig med dets kodificering (jf. fx overskriften “*Flødefod* får det blå stempel” i avisen *Berlingske Tidende*, 13.06.18., 3. sektion, s. 3). DDO synes at have overtaget Dansk Sprognævns (DSN’s) retskrivningsordbogs rolle som ordforrådets symbolske spejl, hvad der er rimeligt nok, ordbogens metodiske grundlag og dens beskrivelsesuniversalitet taget i betragtning (jf. også afsnit 3.5).

3.3. Indtalt udtale

I 2012 blev ordbogen udstyret med en udtalefunktion, der supplerede dens forenkede IPA-transskription, som næppe kan afkodes af alle brugere. Så det må betegnes som et yderligere fremskridt i retning af mere brugervenlighed, hvad fx Bergenholtz & Vrang (2004) har efterlyst ivrigt. Et af skudsmålne var, at DDO er “en ordbog for lingvister”. Det er den *også*, men absolut ikke udelukkende.

Vi gennemførte et par simple tests: bl.a. af minimalparret *mor/mord*, hvor både ordet med og uden stød lød rent og autentisk. Størdordet *hus* kan man desværre ikke få en udtaleprøve på, men derimod på det langt mindre brugte verbum *huse*. Har redaktionen her tænkt på brugere med dansk som modersmål frem for på tilflyttere eller andre fremmedsprogede brugere? Det er rimeligt nok, hvis det er tilfældet, men så bør man være åben om det. Det kan naturligvis også bare være et udtryk for *work in progress*. DDO’s hjemmeside oplyser således, at alle opslagsord “har eller vil blive forsynet med udtale i form af en lydfil”. Det kan man kun se frem til, for også anglicismer som *rockwool*, *windbreaker* og *workaholic* mangler udtalefiler. Der er dog ikke tale om et mønster; ved stikprøverne fandt vi lydfiler i opslagene for adskillige andre anglicismer (fx *frisbee*, *software* og *workshop*). Årsagen kan ligge i ordbogsartiklernes alder – dog var fx *rockwool* allerede med i den trykte udgave af DDO.

3.4. DDO’s app

Appen er med gennemsnitskarakteren 4,1 af 5 stjerner på Google Play blevet ganske godt modtaget af sine brugere, der roser, at den er gratis og reklamefri, samt at den er brugervenlig og tilbyder funktionen *Dagens udtryk*. Dog er der enkelte brugere, der savner grammatiske angivelser, synonymmer (men det var inden funktio-

nen *Ord i nærheden* kom til), stedvist endnu udtalelydfiler og endelig muligheden for at bruge appen offline. Sidstnævnte funktion er dog ifølge DDO bevidst udeladt for at undgå, at ikke-opdaterede udgaver af den online ordbog benyttes.

3.5. Reaktioner på nye ord i DDO

Et interessant og forfriskende indslag ved ordbogssitet er, at det også dokumenterer offentlighedens reaktioner på de nyoptagne ord – næppe bare som en selvbevidst demonstration af ordbogens betydning; formålet er formentlig også at reflektere over interaktionen mellem ordbog og samfund. I hvert fald kommer det til at fungere sådan.

Et spændende eksempel er reaktionerne på behandlingen af adjektivet *grønlænderstiv*: For det første skriver aviserne (ikke mindst de grønlandske), at ordet nu er en “officiel” del af det danske sprog. Det er naturligvis en misforståelse. I Danmark har vi kun en officiel retskrivning, ikke et officielt ordforråd.

Interessant er også både den etymologiske forklaring (“udtrykket bygger på den stereotype opfattelse at grønlændere generelt drikker mere end danskere”) og det ene citat, der formentlig er udvalgt for at tage brodden af ordet: “Udtrykket *grønlænderstiv* er forældet. De seneste 20 år er alkoholforbruget i Grønland nemlig raslet ned, og i dag drikker grønlænderne faktisk mindre end i Danmark”. Dét citat tager de grønlandske medier til sig med kyshånd. Men begge elementer er misvisende. Udtrykket er ikke opstået på baggrund af alkoholstatistikker, men ved selvsyn: Det bygger på erfaringer med i Danmark bosatte grønlandske skæbner, der stikker ud som en markant gruppe af alkoholiserede førtidspensionister – kolonialisme og nationalt ansvar ufortalt. Citatet er også uheldigt, for udtrykket er netop *ikke* forældet. I modsat fald ville det næppe være blevet medtaget i DDO. Distinktionen objekt- og metasprog kan man ikke forvente, at den almindelige

bruger kan håndtere. Bergenholtz & Vrangs (2004) kritik af “en ordbog for lingvister” får altså i dette specifikke punkt en smule medhold.

4. Dobbeltformer og anden varians

Fænomenet “dobbeltformer” i ordbøger er en af DSN anerkendt leksikalsk-ortografisk varians, jf. Galberg Jakobsen (1992). Et eksempel er *for øvrigt* vs. *i øvrigt*. At de er dobbeltformer, betyder, at de to fraser skulle kunne bruges frit udskifteligt med samme betydning. Dobbeltformer, som i princippet kan udarte sig til både trippel- og kvadrupelformer, er et særligt leksikografisk kunstobjekt, der først og fremmest lever i den præskriptive retskrivningsordbog, og måske især i Danmark. Der ville ikke være nogen dobbeltformer, hvis ikke der var en institution, der blander sig i måden, folk skriver på, hvad vi i øvrigt hilser velkomment (jf. Farø 2002). I DDO markeres dobbeltformer ved, at der står to leksemer på opslagspladsen, adskilt med et lille *eller*.

Det kan i praksis og ikke mindst for udenforstående være vanskeligt at skelne mellem dobbeltformer, synonyme, officielle/uofficielle former og anden varians. Og er det nu også virkelig tilfældet, at dobbeltformerne alternerer frit? *For øvrigt / i øvrigt* er måske nok overfladisk betragtet parallelle. Men har de samme brugsegenskaber? Det kan man jo ikke aflæse af leksemernes form, kun via sprogbrugen. Og den har vi mulighed for at analysere vha. det korpus (KorpusDK), som DDO er baseret på. Man kan konstatere, at *i*-varianten forekommer knap 11.000 gange i korpus, mens *for*-formen “kun” optræder knap 600. Det kunne pege på, at *i øvrigt* er normalformen og at *for øvrigt* er den sjældnere, markerede variant.

Balle/balde er et andet, rent ortografisk funderet eksempel: Selv

skriver vi i al uenighed ordet hhv. med *-ll-* og *-ld-*, hvad der kunne være udtryk for forskellige grader af ortografisk konservatisme, jf. staveformen af modalverberne i dansk før retskrivningsreformen i 1948. Ikke desto mindre er formen med dobbelt *-l-* dominerende i sprogbrugen. Hvorfor ikke give i det mindste særligt interesserede brugere denne oplysning, når nu DDO har adgang til den i form af ordbogens egen primærempiri?

Bananasplit/banansplit var indtil for nylig dobbeltformer; tydeligvis var *banansplit* et forsøg på at indføre en dansk form. Men da den ikke slog igennem, trak DSN den tilbage, ligesom tilfældet var med den fortænkte *bruse* som dobbeltform til *bruser*. Det har ført til, at DDO betegner *banansplit* som “nu uofficiel form”. Men hvad med at give oplysningen om, at *banansplit* er en marginal stavevariant? Det er mindst lige så interessant som retskrivningshistorikken – og måske endda praktisk vigtigere.

Handikappet/handicappet er dobbeltformer. Men hvorfor ikke også i samme ombæring nævne synonymet *handikappet menneske*, der er det seneste skud på det politisk korrekte træes stamme? Det ville være nyttigt at få oplysning om forskellige varianter grad af politisk korrekthed.

Dobbeltformen *medium/medie* er et andet eksempel, hvor man kunne ønske sig mere transparens af ordbogen: Det er tydeligt, at der her er tale om et sammenstød mellem en hypernorm og en frekvensnorm. Det er muligt, at DSN vurderer formerne som ligeværdige, men det er de ikke: *Medium* er den konservative, “fine” form (som indtil 1986 var den eneste DSN-godkendte), hvormed man som sprogbruger kan score point i ganske snævre kredse. Vil man derimod skrive, som folk gør flest, skal man bruge den integrerede form, *medie*, som optræder 366 gange i korpus. *Medium* forekommer 282 gange, hvad der kunne ligne en nogenlunde ligelig fordeling. Men 53 af *medium*-eksemplerne er adjektiviske (vedrørende størrelse). Det vil sige, at forholdet er tæt på dobbelt så mange aflatiniserede former som traditionelle. Det kunne godt

formidles til brugeren. Medtager man alle bøjede former, bliver forskellen langt mere markant: 4796 for *medie-* mod som nævnt 282 for *medium*.

Også anden varians, fx i forbindelse med præpositionsvalg, kunne være empirisk mere transparent: Skal man fx kalde sig medlem *af* eller *i* noget (forening, bestyrelse, kirke)? Der kunne man ønske sig en hjælpende hånd på et oplyst grundlag.

Det er som sagt et kernetræk ved DDO, at den er korpusbaseret. Så er det oplagt, at den går linen ud og udnytter denne metodik også der, hvor den nationale ortografiske autoritet (Dansk Sprognævn) fælder sin dom – og afviger fra sprogbrugen. Dette kan man fint referere til, samtidig med at man forbliver sin egen metode tro og “går bag kulisserne” ved at fortælle brugeren, hvad der er op og ned på sprogbrugen. Det viser DDO jo netop i forbindelse med de såkaldte “uofficielle former”, der ikke desto mindre har en stor udbredelse. En sådan *konsekvent empirisme* ville være at foretrække. Også frem for eksempelvis *DanskOrdbogens* såkaldte “proskriptivisme”, hvor baggrunden for de eksplicit anbefalede former heller ikke er transparent nok. Brugeren skal kunne træffe et bevidst valg. Med andre ord: Vil du skrive, som Sprognævnet anbefaler (og dermed signalere sproglig bevidsthed, kulturkonservervisme o.l.), eller vil du skrive som flertallet af danskerne? Denne distinktion er væsentlig og kunne være et fokuspunkt for en kommende strukturel opdatering af DDO.

5. *Tak* – og hvad dertil hører

“Tak er kun et fattigt ord”, hedder en dansk slager fra 1963. Det kunne være nyttigt, om man i en omfattende modersmålsordbog som DDO kunne få oplysninger om den slags bevingede ord, der er gået ind i sprogets ord- eller tekstforråd. Modstanden mod det encyklopædiske stof i ordbøger er velkendt og ofte blevet tilba-

gevist som unødvendig (jf. fx Farø & Hyllested 2007). I det omfang, bevingede ord har fået etablerede semantiske og pragmatiske funktioner, er de højrelevant stof i en monolingval ordbog. Ikke mindst fordi man typisk ikke finder sproglige oplysninger i opslagsværker over bevingede ord (se fx Jarvad 2006 og Büchmann 2007), der hovedsageligt interesserer sig for det encyklopædiske, selv om ordbøgerne heller ikke på dette område leverer tilfredsstillende med information.

Nu er det ikke bevingede ord, der er dette afsnits hovedfokus, men derimod talehandlingen *tak*. For det første er der selve substantivet *tak* og de mange præpositionalobjekter, der sammen med kernesubstantivet udgør etablerede flerordsforbindelser, først og fremmest *tak for* + NOM: *tak for alt / besøget / dansen / denne gang / det / en dejlig aften / gaven / hjælpen / i aften / i dag / i går / invitationen / kaffe / kampen / lån / mad / sidst / skænken / snakken / tippet / turen* m.fl.

Andre vigtige forbindelser med *tak* er de kausale: *tak fordi du/I kom / fordi du/I ville være med*. Endvidere kan man kombinere *tak* med præpositionen *til* for at tilknytte en adressat: "Tak til Farmor!". Men der er også disse varianter: *selv tak / tak selv / tak som byder / tak i lige måde / tak skal du/I have*.

Dette (ufuldstændige) inventar er resultatet af en brainstorm. Den burde suppleres af en grundig korpusresearch. Men hvordan ser DDO's behandling ud? Tilsyneladende er hovedfokus på reception. Det må være konklusionen på, at den vælger at nævne *eksempler* på kombinationer med *tak* i stedet for at opregne det samlede leksikalske inventar af takkeformler.

Det er ikke indlysende, at *tak for mad, te* etc. behandles som sublemmaer, mens *tak for lån, turen, i går og kampen* blot er eksempler på brugen af *tak*. I øvrigt har *tak for kampen* også en idiomatisk betydning, nemlig spøgende taknemmelighed for et netop overstået samarbejde. *Tak for dansen* og formlens idiomatiske betydning ('tak for samarbejdet' og som vittig kommentar, når koor-

dinationen mellem to mennesker, der prøver at gå forbi hinanden, mislykkes), kunne også nævnes.

Ordbogen nøjes med at angive de deciderede idiomatiske fraser som *tak for kaffe*, der som idiom (og med særegen intonation) er udtryk for stor forbavselse, ofte i negativ retning. Men det er vigtigere at beherske de kulturelt væsentlige og højfrekvente faste forbindelser (jf. *rutineformler* i Coulmas' 1981 terminologi) end at forstå betydningen af et ikke voldsomt hyppigt idiom. Derfor kunne ordbogen i højere grad end nu prøve at sætte fokus på det leksikaliserede inventar af takkeformler frem for blot at nævne eksempler.

6. Neologismer og andre potentielle opslagsord

Vi har foretaget en brainstorm med målet at ramme især nyere ord og udtryk, der burde tjekkes i ordbogen. Vi har i det hele taget ladet fantasien løbe, og tabel 1 gengiver et udvalg af dem af vores indfald, der enten ikke er optaget eller tilfredsstillende dækket i DDO (en del af de undersøgte ord fandtes dog i DDO, fx *blogger*, *gamer*, *hijab* og *vegansk*). Desuden vises det antal træffere, en Google-søgning på udtrykkene gav.³ Vi er bevidste om, at disse tal ingen statistisk værdi har, men de giver dog en indikation om ordenes og udtrykkenes udbredelse – også i den aktuelle, ikke-normerede sprogbrug, som Google, i modsætning til tekstkorpora, viser. Antallet af træffere er i nogle tilfælde angivet lavere, end det formentlig er. Det skyldes, at leksemerne blev søgt med danske bøjningsendelser (fx præteritumformen *tweetede*) og sammen med ord, der er eksklusive for dansk, med eller uden citationstegn, fx “stor koalition”, “store koalition”. Hyppigheden for udtryk som *controller*, *dø* og *poste* (i de betydninger, der mangler i DDO) samt *åh-åh* var

3 Ved ord og udtryk, der findes i DDO, men ikke er dækkende behandlet, er hyppigheden fundet med Google ikke angivet.

det dog ikke muligt at danne sig et billede af med Googles hjælp, da de mange tusinde resultater ville skulle vurderes enkeltvist.

<i>amagergevir</i> ; 0 (men <i>røvgevir</i>)	<i>amagermåge</i> 0
<i>amagernummerplade</i> 700	<i>avatar</i> 427.000
<i>benmølle</i> 740	<i>brøndbyfontæne</i> 0
<i>chicagosokker</i> 0	<i>CO₂-forbrug</i> 9.650 (men <i>-afgift</i>)
<i>controller</i> (mgl. gamer-bet. 'joystick')	<i>damper</i> (e-ryger) 194.000
<i>dirndl</i> 23.500	<i>dø</i> (mgl. gamer-bet. 'tabe' spillet / et liv)
<i>eid</i> 399.000	<i>elevintra</i> 244.000
<i>fake news</i> 326.000 (fakemedier)	<i>footies</i> 89.400
<i>grindedrab</i> 26.100	<i>grungeskæg</i> 3
<i>hadtale</i> 23.900	<i>hiposvin</i> (uden betydningsangivelse)
<i>hipsterskæg</i> 14.200	<i>hirtshalshabit</i> 447
<i>influenzavaccination</i> 307.000	<i>iPad(-)skuldre</i> 200
<i>ishøjpalme</i> 272	<i>jordkoger</i> 21
<i>medlemsapp</i> 26.900	<i>middelaldermarked</i> 52.700
<i>myrelokkedåse</i> 4.800	<i>nexøhabit</i> 48
<i>palæo-</i> mgl. bet. 'grov' (om mad)	<i>plastikaffald</i> 17.800
<i>poste</i> mgl. bet. 'slå op på sociale medier'	<i>prank</i> (practical joke) 265.000
<i>raw-</i> (-kage, -opskrift m.m.)	<i>rawfood</i> 94.600
<i>slowjuicer</i> 41.700	<i>smartphone(-)nakke</i> 265
<i>sojakød</i> 2.230	<i>stenalder-</i> 20.000 (brød, mad, m.m.)
<i>stor koalition</i> 15.000	<i>støvsugerskæg</i> 0
<i>svinebæst</i> 92	<i>talesyntese</i> 49.100
<i>to go</i> 67.300 (men <i>to go-krus</i>)	<i>tweeter</i> 1.500
<i>våbenlobby</i> 27.100	<i>youtuber</i> 143.000

Tabel 1: Ord og udtryk fra artiklens forfatteres brainstorm med antal forekomster på Google.

Det er klart, at nogle af vores indfald er vildskud: Når et ord har o træffere på internettet, som det var tilfældet med *amagermåge*, *-gevir*, *brøndbyfontæne*, *chicagosokker* og *støvsugerskæg*, så er det nok ikke udbredt i dansk – om end der kunne være tale om særmundtlige former – eller snarere idiolektale ad hoc-dannelser. For at kunne associeres med dansk ordforråd må der et vist antal belæg i forskellige tekster af forskellige sprogbrugere til. Vi forestiller os et minimum på omkring 100 belæg som delkriterium, hvad der fx vil udelukke *nexøhabit* og *grungeskæg*. Til gengæld kunne man overveje at oprette sidsteleddet *-habit*, der i sine forskellige kombinationer tilsammen er temmelig produktivt. I det hele taget kunne ordbogen have større fokus på produktive bundne morfemer og morfembetydninger, jf. fx *raw-*, *-liderlig* ('betaget af'), *to go*, *palæo-* og *stenalder-* ('grov-').

7. Tekstbaserede stikprøveundersøgelser

Vi benytter os her af metoden tekstbaseret stikprøveundersøgelse. Stikprøverne er foretaget i to ret forskellige medier, nemlig i *Weekendavisen* (8 og 9, 2018) og *Anders And* (70:5, 2018). Vi ville undersøge, om den tekstbaserede kontrol af lemmabestanden var mere eller mindre effektiv end den helt frie. I begge metoder spiller intuition en væsentlig rolle, for uden den ville det for en menneskelig kontrollant være et uoverskueligt arbejde. Intuitionen sørger for en grovsortering blandt de tusindvis af leksikalske kandidater, der trænger sig på.

7.1. Weekendavisen 2018, nr. 8 og 9

Vi lægger ud med en undersøgelse baseret på *Weekendavisen*. Ligesom ved brainstormen fandt vi en del af de undersøgte ord og udtryk i DDO, fx *celledeling*, *kvadrant*, *statsbudget* og *urbanisering*.

Tabel 2 viser ord og udtryk, der ikke var i DDO *eller* ikke var dækkende behandlet, samt antal Google-hits i dansk kontekst (afrundet).

<i>Darfur-regionen</i> 100	<i>dokumentationskrav</i> 83.400
<i>evolutionsbiologi</i> 5.110	<i>FN-rapport</i> 2.660.000
<i>formidlingsevne</i> 65.700	<i>forskerhold</i> 67.100
<i>forudprogrammeret</i> 10.200	<i>genaktualisere</i> 1050
<i>innovationsminister</i> 152.000	<i>kontekstafhængig</i> 32.800
<i>kontekstuaafhængig</i> 2.180	<i>maskeringsforbud</i> 205.000
<i>populationsgenetik</i> 4.650	<i>robotteknologi</i> 98.500
<i>singularist</i> 2.000	<i>sprogpurist</i> 320
<i>ultraliberal</i> 23.900	<i>våbenforherligelse</i> 3
<i>våbenlovgivning</i> 11.600	

Tabel 2: Ord og udtryk fra *Weekendavisen* 2018, nr. 8 og 9.

7.2. Anders And 2018, nr. 5

Efter samme princip som ved de to foregående tabeller vises i tabel 3 de ord og udtryk, der ikke er inkluderet i DDO. De fleste undersøgte ord fandtes dog i DDO, deriblandt *benmuskler*, *gå løs*, *tredjepremie* og *tøset*.

<i>curlingsten</i> 7.950	<i>giftpil</i> 9.090
<i>hus</i> (curling)	<i>hvem prøver du/I at narre/snyde?</i> 500 ⁴
<i>(komme) op på hesten igen</i> 87.000	<i>timetakst</i> 11.600
<i>vikingefestival</i> 26.500	<i>åh-åh</i> . Note: første stavelse kort med stød, den næste længere og uden stød, tonelejet er henholdsvis højt og dybt. Interjektionen (eller udtrykspartiklen) bruges som reaktion på en uheldig situation.

Tabel 3: Ord og udtryk fra *Anders And* 2018 nr. 5.

4 Inklusive kløvede former: *Hvem er det, du/I prøver at narre/snyde?*

8. Delkonklusioner og diskussion

Hvis vi sammenligner de to metoder til lemmakontrol, brainstorm og tekstbasering, så er de begge udmærkede til at detektere lakuner med. Der var dog ikke så mange huller på basis af Anders And-bladet, hvilket kunne skyldes, at netop denne kilde – nok i høj grad af hensyn til sit kernepublikum – holder sig til det basale og til dels det centrale ordforråd. De tre empiriske konfrontationsundersøgelser demonstrerer derudover følgende: DDO er relativt godt dækket ind mht. opslagsord; vi finder en stor del af det, vi søger, både det intuitivt søgte og det tekstuelts udsprungne “opslagsbehov”. Men der er stadig en del mangler, hvad der ikke overrasker os og næppe heller redaktionen selv; det vil der blive ved med at være.

Vi er både stødt ind i meget udbredte og en række relativt lavfrekvente, men tilsyneladende etablerede og også forklaringskrævende ord og udtryk, der lige så godt kunne optages og beskrives af ordbogen. Hvorfor? Fordi en korpusbaseret almenordbog ikke nødvendigvis kun skal dække det basale og det centrale ordforråd. Dette *skal* selvfølgelig dækkes. Men hvad med det mere lavfrekvente og dermed tilsyneladende perifere, men ikke desto mindre etablerede? Hvis det basale og centrale efterhånden er beskrevet godt og solidt, er der ikke noget i vejen for, at også mindre hyppige udtryk kan komme med, navnlig de særligt prægnante eller semantisk uigennemskuelige. Naturligvis giver det ikke mening at beskrive stærkt faglige, idiolektale og klart uetablerede ord og udtryk. Men en ordbog, der er online, behøver ikke holde sig tilbage. Det eneste, denne nye tilgang til ordforrådet ville kræve, var, at ordbogen markerede sit indhold som hhv. høj-, mellem- og lavfrekvent; gerne grafisk, det er mere pædagogisk og for de fleste brugere fuldt tilstrækkeligt. Det er der også præcedens for, eksempelvis i MED og *Duden online*. Og det må være en korpusbaseret online-ordbogs fornemste opgave frem for at fokusere på at definere en

nedre frekventuel grænse for relevans. Det afgørende i dette stadium af projektet, hvor ordbogens karakter af dokumentationsordbog af en bestemt periodes ordforråd (som ved den trykte ordbog) er forladt, må være et andet og tilsvarende ambitiøst mål, nemlig at brugeren kan få en forklaring på det, vedkommende måtte slå op. Gerne også det, der rækker ud over det centrale ordforråd. For netop det repræsenterer for mange det vanskelige og opslagskrævende.

9. Sammenfatning

DDO vokser løbende, og det endda i et relativt stærkt tempo. Det er glædeligt, og det er i sig selv grund nok til at motivere online-versionen ved siden af det respektindgydende seksbindsværk. Uden online-versionen ville opdatering være en meget krævende opgave. Så version 2.0 skabte grundlaget for ordbogens fortsatte levedygtighed og dermed indirekte 3.0, og tak for det, siger vi brugere, der opdrager studerende til at bruge både ordbog og korpus. Men der er (heldigvis?) masser at gøre på ordbogen endnu.

Vi har konstateret, at der stadig er huller, både sådanne, der vedrører det nyttilkomne ordforråd, som det kan være vanskeligt at være 100 % opdateret på, men også på områder, hvor leksemerne længe har været etablerede. Herudover foreslår vi, at ordbogen udvider sin genstand til også at omfatte leksikaliserede elementer, der måske har en lavere frekvens end opslagsværkets nuværende kriterier for inklusion, men som af andre grunde, fx ugennemskuelighed, prægnans o.a. kan have brugernes interesse. Endelig har vi foreslået, at bevingede ord vies systematisk opmærksomhed. Navne er en anden mulig kategori – se hertil Trap-Jensen & Lorentzen (2017).

På den strukturelle side har vi argumenteret for en større autonomi i forhold til Sprognævnet, således at dette organs holdninger

ikke kommer på tværs af DDO's konception og metode: DSN's beslutninger kan udmærket gengives, der hvor nævnet har autoritet til det. Men det bør ikke føre til, at det bliver en sovepude i forhold til ordbogens primære funktion: At beskrive det danske sprog, som det bruges. Derfor kunne DDO også udnytte korpusmetoden konsekvent der, hvor DSN har erklæret valgfrihed. Dets beslutninger har ofte måttet trækkes tilbage, fordi den skrivende befolkning ganske enkelt ignorerer dem. Her kunne DSN muligvis lære af DDO, så det undgik for mange urealistiske fejlskud, der jo i sidste ende svækker nævnets autoritet. Den fare udsætter DDO sig ikke for, så længe den er konsekvent i sin metode og i sin kommunikation af den til kolleger og offentlighed.

Litteratur og kilder

Ordbøger

Büchmann, Georg (2007): *Der Große Büchmann. Geflügelte Worte von Aristoteles bis Zappa*. Berlin: Knauer.

DanskOrdbogen (1999): Henning Bergenholtz & Handelshøjskolen i Århus, Center for Leksikografi (1. udgave, 1. oplag). Århus: Center for Leksikografi.

DDB (2014) = Nimb, Sanni, Henrik Lorentzen, Liisa Theilgaard & Thomas Troelsgård (2014): *Den Danske Begrebsordbog* (1. udgave). Odense: Syddansk Universitetsforlag.

DDO (2009-) = Det Danske Sprog- og Litteraturselskab: *Den Danske Ordbog* (online). <<https://ordnet.dk/ddo>> (juni 2018).

Duden online. <www.duden.de> (juni 2018).

Duden-ordbøgerne (1996) = Dudenredaktion & Scholze-Stubenrecht, Werner: *Duden, Rechtschreibung der deutschen Sprache* (21. völlig neu bearb. und erw. Aufl. ed.). Der Duden in 12 Bänden, 1. Mannheim: Dudenverlag.

- Jarvad, Pia (2006): *Bevingede ord*. 7. udgave. Fuldstændig omskrivning og nyrevidering af T. Vogel-Jørgensens 6. udgave. København: Gads Forlag.
- MED (2007) = *Macmillan English Dictionary for Advanced Learners* (2nd ed.) Oxford: Macmillan Education.
- Nudansk Ordbog* (2010). København: Politiken.
- ODS (1918-1956) = *Ordbog over det danske Sprog*. Det Danske Sprog- og Litteraturselskab. København: Gyldendal.

Anden litteratur og empiri

- Anders And*. Egmont (2018), 70. årgang, nr. 5.
- Bergenholtz, Henning & Vibeke Vrang (2004): Den Danske Ordbog imponerer og skuffer. I: *Hermes* 33, 149-178.
- Berlingske Tidende*, 17.06.18, 3. sektion, 3.
- Coulmas, Florian (1981): *Routine im Gespräch. Zur pragmatischen Fundierung der Idiomatik*. Wiesbaden: Akademische Verlagsgesellschaft Athenaion.
- DSL (2018) = Det Danske Sprog- og Litteraturselskab (2018): *Forrygende eller skæppeskønt? Nyt i Den Danske Ordbog*. <dsl.dk/presse/dokumenter/pressemeddelelse-nyt-i-den-danske-ordbog> (juni 2018).
- Farø, Ken (2002): Ortografi og empiri: konsonantfordobling eller ej? I: *LEDA-nyt* 34, 4-9.
- Farø, Ken (2018): Tak for alt! Om takkeformler i dansk og tysk. I: Inger Schoonderbeek Hansen, Kirsten Lyshøj & Viggø Sørensen (red.): *Fraseologi – genveje og omveje. Festskrift til Torben Arboe*. Aarhus Universitet, 51-56.
- Farø, Ken & Adam Hyllested (2007): Ordbog, leksikon eller encyklopædi? Et koncept for et dansk lingvistisk opslagsværk. I: *LexicoNordica* 14, 107-144.
- Galberg Jakobsen, Henrik (1992): *Danske dobbeltformer: valgfri former i retskrivningen* (Dansk Sprognævns skrifter 18). København: Munksgaard.

- Google Play: Brugeranmeldelser af appen Den Danske Ordbog. <play.google.com/store/apps/details?id=dk.dsl.ordnet.ddo &hl=da> (juni 2018).
- KorpusDK. Det Danske Sprog- og Litteraturselskab. <https://ordnet.dk/korpusdk> (juli 2018).
- Leroyer, Patrick (2005): Ordbogseksterne kildehenvisninger – fra belæg til funktioner. I: Henrik Lorentzen & Lars Trap-Jensen (red.): *Nordiske Studier i Leksikografi* 8. København: Nordisk Forening for Leksikografi, 239-251.
- Pálfi, Loránd-Levente (2005): Den Danske Ordbog – en kritisk opsamling af den offentlige ordstrid. I: *LexicoNordica* 12, 123-149.
- Trap-Jensen, Lars & Henrik Lorentzen (2017): Navnestof i ordbøger – nye tider, nye måder. I: *LexicoNordica* 24, 109-129.
- Weekendavisen* (2018) nr. 8 og 9 (uddrag).

Ken Farø
lektor, ph.d.
Københavns Universitet
Institut for Engelsk, Germansk og
Romansk
Emil Holms Kanal 6
DK-2300 København S

Lene Vinther Jensen
akademisk medarbejder, cand.mag.
Københavns Universitet
Institut for Engelsk, Germansk og
Romansk
Emil Holms Kanal 6
DK-2300 København S