

Fra historisk værk til moderne onlineresurse: NAOB som netordbog

1 Indledning

Da *Det Norske Akademis ordbok* (NAOB) gik i luften den 27. december 2017, fik den norske offentlighed fri adgang til en omfattende videnskabelig betydningsordbog for den variant af norsk som størstedelen af nordmændene bruger. Datoen markerer samtidig afslutningen på et projekt der har strakt sig over 20 år og har forvandlet den historiske papiordbog *Norsk riksmålsordbok* (NRO) til en moderne, opdateret netordbog for bokmål. Det er naturligvis en historisk begivenhed i norsk leksikografi, men også en markant begivenhed i nordisk sammenhæng. I sit anlæg og omfang er NAOB en nationalordbog i traditionen fra *Svenska Akademiens ordbok* (SAOB) i Sverige, *Ordbog over det danske Sprog* (ODS) i Danmark og Norges anden nationalordbog, *Norsk Ordbok*, for nynorsk. Antallet af opslagsord svarer stort set til antallet i ODS (ca. 225.000).

I denne artikel vil jeg først og fremmest vurdere NAOB som en netordbog for moderne brugere. Som ikke-norsktalende har jeg i sagens natur en begrænset viden om og indsigt i de helt sprognære beslutninger redaktionen har skullet træffe, herunder ikke mindst de valg af morfologiske varianter (opslagsformer, varianter og bøjningsformer) som har skullet træffes for at omdanne en ordbog over riksmålet til en ordbog for moderat bokmål. Mit kendskab til norsk sprogdebat er til gengæld stort nok til at jeg afholder mig fra at have en mening om dette aspekt og trygt overlader det til andre og mere kyndige at kommentere den side af NAOB.

Selvom hovedvægten vil være på netordbogen, vil jeg dog også kort berøre de øvrige elementer der indgår i transformationsprocessen fra papir til skærm. De enkelte delelementer er indbyrdes afhængige på den måde at valg der træffes ét sted i processen, sætter rammerne for hvilke løsninger der er mulige under de øvrige elementer. Lad os derfor først

opdele processen i en række faser og se på hvordan de er prioriteret og løst.

2 Fra papirordbog til moderne netordbog

Første fase i processen er retrodigitaliseringen: omsætningen fra traditionel bogside til et digitalt format der kan fungere som udgangspunkt for skærmvisning. Her har redaktionen skullet vælge mellem OCR-scanning og indtastning som digitaliseringsmodel og har valgt scanningen som løsning. Med den kvalitet som OCR-scanning har med den nuværende teknologi, viser det sig ofte at være en dyrere løsning end indtastningsmodellen. Retrodigitaliseringen af ODS skete således ved hjælp af dobbeltindtastning (se Bojsen og Trap-Jensen 2005), og også SAOB har valgt at lade hele manuskriptet gendigitalisere efter ODS-modellen (se Bäckerud 2014) efter at den oprindelige scanningsløsning (fra 1987) viste sig at indeholde for mange fejl. Problemet med scanning er normalt at en efterfølgende korrekturlæsning af hele manuskriptet er nødvendig, og udgifterne hertil kan hurtigt overstige den besparelse som den billigere scanningsløsning udgør i udgangspunktet. I NAOB's tilfælde er valget af scanning dog forståeligt eftersom langt de fleste artikler alligevel har skullet bearbejdes redaktionelt, og i den forbindelse har korrekturlæsning af scanningsresultatet været en overkommelig opgave. Når man betragter resultatet, er det da også lykkedes på imponerende vis, og jeg er endnu ikke stødt på scannefejl, hvilket kan ses som en bekræftelse på at redaktionen har tjekket og rettet hver enkelt artikel. Det er nydeligt gjort.

Næste fase i processen består i at parse den digitaliserede tekst, dvs. fortolke manuskriptets typografiske markeringer og omsætte dem til indholdselementer. Konverteringen kan kun udføres som en delvis automatiseret proces og foregår typisk trinvis. Det er eksempelvis ret nemt at identificere ord skrevet med fed skrift og med indrykning efter linjeskift som lemmaets opslagsform og tekst i kursiv efterfulgt af en ordinær parentes som citat + kildeangivelse. Definitioner optræder med ordinær skrift, ofte umiddelbart efter en angivelse af betydningsniveau anført med romertal, arabertal samt bogstaver i enten fed eller ordinær – men definitionerne er ikke den eneste oplysningstype der er angivet på denne måde. Konverteringen forløber altså som en gentagen proces, idet stadig flere indholdselementer genkendes og isoleres indtil der bliver en rest tilbage som kræver manuel stillingtagen. Hvor langt man kan komme ad

automatisk vej, afhænger ikke mindst af hvor stringent redaktørerne har fulgt principperne for ordbogen i det oprindelige værk, og hvor meget redaktionen ønsker at ændre den oprindelige struktur så den passer til en moderne ordbog. Den opmærksomme bruger indser hurtigt at NAOB har gjort meget for at forenkle artikelstrukturen så den er fladere: De nummererede betydninger holder sig, så vidt jeg kan bedømme, til højst fire niveauer (se f.eks. artiklerne *etter*, *en*, *den*), og NRO's brug af romertal og bogstaver er helt sløjfet (omend opslagsord godt kan rumme flere ordklasser, fx præposition + adverbium eller determinativ + pronomen). Langt de fleste artikler kan dog nøjes med to niveauer: hovedbetydning og underbetydning. Det giver en overskuelig og brugervenlig artikelstruktur. At nå dertil har krævet en betydelig redaktionel indsats hvor redaktørerne har skullet tage stilling til om betydninger på forskellige niveauer i NRO har kunnet slås sammen til én. Derved går naturligvis nogle nuancer tabt, men det er den pris man må betale for at få øget overskuelighed. Jeg synes det fungerer fint.

Faktisk kunne jeg ønske mig at redaktionen på tilsvarende vis havde udstrakt forenklingen til også at omfatte flerordsudtryk. På det punkt har redaktionen bevaret NRO's placering af flerordsudtryk under den betydning som udtrykket må henregnes til. Hvis udtrykkene flyttes fra den oprindelige betydning til én samlet blok, går der naturligvis også nuancer tabt – ligesom det er tilfældet når flere betydninger slås sammen. Til gengæld koster det noget i overskuelighed og brugervenlighed at bevare dem under en bestemt betydning. Lad os se på et eksempel.

Hvis man er i tvivl om betydningen af udtrykket *se ut* og slår op i NAOB for at få hjælp, må man orientere sig op til ti forskellige steder i artiklen *se* for at finde den relevante forklaring. Den kan befinde sig under en betydning af de respektive udtryk *bli seende ut som*, *ikke se ut*, *se seg ut* eller *se ut som*. Det kræver altså en vis tålmodighed af brugeren at finde frem til den rette beskrivelse. Artiklerne er ganske vist blevet forsynet med en indledende indholdsfortegnelse – for udtrykkenes vedkommende kaldt en *uttryksfortegnelse* – og disse hjælper nok til at skabe overblik og hurtigere adgang, men trods det prisværdige initiativ siger erfaring fra brugerundersøgelser af *Den Danske Ordbog* (DDO) at brugerne desværre sjældent benytter sig af overblikvisningen (Svendsen 2016: 60). Hvis man virkelig ønsker at hjælpe brugerne, er det efter min mening derfor bedre at samle de faste udtryk i én blok efter de nummererede betydninger. De fleste brugere læser alligevel ikke en stor artikel

i sin helhed, og dermed er der ikke tungtvejende argumenter for at bevare placeringen af dem under en bestemt betydning. Redaktionen har muligvis overvejet dette, men valgt at prioritere anderledes, for ændringen er ikke alene gjort ved at flytte udtrykkene fra ét sted til et andet. Undertiden kræves tillige en redaktionel indsats, for eksempel hvis en betydningsdefinition er afhængig af en nedarvet beskrivelse fra et højere niveau. I figur 1 kan citatet “*jeg vet fan hvorledes det er*” ikke nemt forstås hvis det fjernes fra sin overbetydning 1.2 ‘i eder og kraftuttrykk, til dels som interjeksjon’ og måtte i det hele taget gives en egen lemmaform, f.eks. *vite fan hvorledes noe er*. Det vil sige at hvis udtrykkene flyttes, kræver det at udtrykket forsynes med en selvstændig og uafhængig definition. Men også dette er vel brugervenligt.

1.2.3 BRUK I NEKTENDE UTTRYKK | jf. [heller](#)

SITAT

- *jeg vet fan hvorledes det er* (Gunnar Heiberg: *Samlede dramatiske verker* II 95)

UTTRYKK

du skal faen heller

det skal du absolutt ikke

ikke faen

- *det gjør jeg ikke! Ikke faen!* (Agnar Mykle: *Lasso rundt fru Luna* 112)

1.2.4 underfundigheter; fandenskap

UTTRYKK

være full av fanden/faen

- *hun skulde ha vært i fengsel, og ... hun var full av fa'en* (Dagbladet 1927/219/3/4)
- *- Han var full av faen, og det var godt å bli kvitt ham, skål for det* (Bernt Vestre: *En hane til Asklepios* 58)

Figur 1. Betydningsforklaringer kan nedarves fra den betydning udtrykket står under.

En yderligere vanskelighed med at finde et bestemt udtryk er udtrykkenes rækkefølge. Udtrykkene er nemlig ikke (altid) alfabetiserede i artiklerne, hvilket ikke gør det nemmere for brugeren at finde et bestemt udtryk; man kan faktisk blive nødt til at se hele artiklen igennem for at finde et konkret udtryk. Figur 2 viser et eksempel fra artiklen *hjerte*, hvor udtryk nævnt under betydning 2.1 opregnes i den rækkefølge de optræder i, dels i NAOB, dels i NRO.

De enkelte udtryk i NAOB er blevet bearbejdet og revideret redaktionelt i forhold til NRO, både med hensyn til lemmaform og ortografi, og endvidere ved at rækkefølgen er blevet ændret. Men altså ikke til et alfabetisk ordningsprincip, og vejledningsteksten er tavs om dette forhold.

NAOB	NRO
<i>av hjertets/hjertens fylde</i>	<i>det er som talt ut av mitt hjerte</i>
<i>av hjertens grunn</i>	<i>av sitt hjertes fylde, av hjertets (hjertens) fylde</i>
<i>av et godt hjerte</i>	<i>av hjertets grunn</i>
<i>av hele ens hjerte</i>	<i>hate en av et godt hjerte</i>
<i>i sitt hjerte</i>	<i>av hjertet</i>
<i>være som talt ut av ens hjerte</i>	<i>etter ens hjerte</i>
<i>av hjertet</i>	<i>komme fra hjertet</i>
<i>(helt) etter ens hjerte</i>	<i>i sitt hjerte</i>
<i>komme fra hjertet</i>	

Figur 2: Udtryk nævnt under betydning 2.1 af hjerte 'LITTERÆRT (ens) innerste, dypeste sinn eller følelse' i NAOB og NRO.

Til gengæld er udtrykkene alfabetiseret i den indledende udtryksfortegnelse. Det kunne tyde på at der er tale om et valg, men jeg er ikke i stand til at gennemskue hvad det ordnende princip er.

Et andet argument for at samle de faste udtryk i én blok, er at man så ikke behøver at sende brugeren unødigt rundt i værket ved hjælp af henvisninger. I eksemplet *se ut* optræder flere af udtrykkene også i artiklen *ut*, med henvisninger til *se*. Det er dog ikke gjort konsekvent; der er eksempelvis ikke henvisning fra *seg* til *se seg ut* (under *se*), hvorimod der henvises fra *seg* til *ta seg ut* (under *ta*). Henvisningerne er jo et levn fra den trykte bog, mens udtrykkene i en onlineordbog fint kan stå under hvert af de indgående ord hvis de frigøres fra en bestemt betydning. På den måde ville brugeren slippe for at blive sendt rundt i ordbogen. Og udtrykkene behøver naturligvis ikke optræde mere end én gang i databasen; det kan klares i det stilark der beskriver hvordan indholdet præsenteres på netsiden. Forudsætningen er blot at de enkelte ord i et givet udtryk er opmærket med oplysning om hvilket lemma de tilhører.

Skulle redaktionen prioritere denne opgave, vil det have den yderligere fordel at de faste udtryk bliver gennemset systematisk. Med den nuværende visning er det dels vanskelig for brugerne at vide hvor de kan finde et bestemt udtryk, dels finder man overflødige dobbeltbeskrivelser af udtryk, enten overtaget direkte fra den trykte ordbog eller redigeret uafhængigt af hinanden. Hvis man eksempelvis betragter udtryk placeret under *stor*, er det vanskelig at gennemskue hvorfor *stor bakke* og *stort intervall* forklares under *stor*, mens *stor bokstav* og *stort gross* henviser brugeren til *bokstav* og *gross*. Udtrykket *gjøre store øyne* forklares til gengæld

både i artiklen *stor* ('i udtrykk for forundring, redsel e.l.') og i artiklen *øye* (med en lidt anden forklaring: 'sperre øynene opp av forbauselse'). Andre eksempler på dobbeltbeskrivelse jeg er stødt på, er *gå gjennom marg og ben* (under *ben* og *marg*), *fet betong* (under *betong* og *fet*), *for mye av det gode* (under *god* og *mye*), *som man står og går i* (under *gå* og *stå*) og *gå stille i dørene* (under *dør* og *stille*). Nogle gange er udtrykkene redigeret uafhængigt af hinanden, andre gange står samme beskrivelse i begge artikler.

Og hermed har vi allerede bevæget os et stykke ind i næste fase.

3 Indholdsredigering og opdatering af NRO's eksisterende artikelbestand

Hvor ændringer og tilpasning af datastrukturen til den moderne ordbog er en delvis automatiseret proces, er indholdsredigeringen en manuel og resursekrævende indsats. Her er gjort et stort og fint arbejde med at gennemrevidere artikelbestanden så den fremstår tidssvarende og beskrevet i et ligefremt og letforståeligt sprog for moderne brugere.

Selvom denne fase er principielt uafhængig af den foregående, er arbejdet i praksis overlappende. Når betydninger slås sammen for at opnå en fladere artikelstruktur, indebærer det jo at der skal udformes en definition der omfatter flere tidligere definitioner, dvs. at strukturelt begrundede beslutninger får konsekvenser for den redaktionelle revision. Nilstun (2012) beskriver hvordan NRO's definitioner ofte bærer præg af for lavt abstraktionsniveau og for høj grad af encyklopædisk præcision, dvs. at definitioner er udformet så de passer på enkeltcitater fra excerpt-samlingen, og at en del af arbejdet med NAOB har bestået i at udforme mere generelle definitioner når betydningerne slås sammen til en grovere betydningsstruktur. Nilstun udtrykker selv håb om at resultatet vil være "balansert når det gjelder disse to motpolene" (Nilstun 2012: 127). For betydningsopdelingen synes jeg det er lykkedes fint, men for definitionernes vedkommende er de undertiden vel kortfattede. Især adjektivdefinitioner består ofte af blot et enkelt ord, et synonym. Det kan undertiden være på sin plads, men som hovedregel kun hvis opslagsordet er sjældent eller vanskeligt og synonymet er mere almindeligt. Det er ikke altid tilfældet i NAOB, som når *veldig* forklares som 'mektig' (bet. 1), *mektig* som 'virkningsfull' (bet. 2.1) og *sterk* som 'tallrik' (bet. 1.3). Især

er det problematisk hvis det forklarende ord i sig selv er polysemt. Det hjælper dog hvis ordet er klikbart, således at den betydning der er mål for linket, har en diskret kasse omkring sig, og det er heldigvis ofte tilfældet.

NAOB indeholder også eksempler på cirkelforklaringer af typen *lei* = 'kjed; trett', *kjed* = 'lei; trett', *trett* = 'lei; kjed' eller *betrakte som* = 'holde for, regne for', *regne for*, *regne som* = 'anse som', *anse som/for* = 'betrakte som, regne som', *holde for* = 'anse, regne for'. Det er næppe et stort problem, men vel mere alvorligt når enkeltordsdefinitioner er så forholdsvis udbredte. Det er noget som redaktionen kan arbejde på at revidere. Definitionstypen med et eller to synonymer gør at NAOB nok ikke er så enkel at bruge for den der ikke har norsk som modersmål.

Når det gælder indholdsredigering, har reaktionen gjort et stort arbejde for at supplere lemmalisten med mange ord fra de seneste årtier, således at ordbogen kan leve op til prædiketet som diakron samtidsordbog. Det er ikke lykkedes mig at finde dokumentation for antallet af ny-skrevne artikler, men det generelle indtryk er at redaktionen har været flittig og prioriteret dette arbejde højt ved at tilføje mange ord fra det moderne ordforråd, herunder ikke mindst ungdomssproget fra eksempelvis serien *Skam*. Her finder man eksempler som *disse*, *ditche*, *fuckboy*, *lættis*, *random*, *serr*, *sykt* og mange flere. Det er glimrende. Til gengæld har redaktionen tilsyneladende også fjernet en del ord, og det er mere overraskende. Heller ikke her har jeg kunnet finde dokumentation for omfanget, men uden at anstrenge mig kan jeg ved at sammenligne NRO og NAOB hurtigt finde adskillige eksempler: *børsteindustri*, *continuetur*, *efeserbrev*, *enromshus*, *enromsstall*, *gotråd*, *hjemmefarvann*, *hjemmeforbruk*, *mestvinnende*, *preleksjonskatalog*, *sosse* (og nogle sammensætninger: *sossegjeng*, *sossejente*, men ikke *sossested*, som er med) er alle ord der står i NRO's bind 5 og 6, men ikke i NAOB. Og et ord som *EF* og samtlige sammensætninger hermed er systematisk udeladt: *EF-avstemning*, *EF-kamp*, *EF-kampanje*, *EF-kommisjon*, *EF-land*, *EF-medlemskap*, *EF-motstand*, *EF-motstander*, *EF-område*, *EF-oppgjør*, *EF-parlament*, *EF-sak*, *EF-soner*, *EF-spørsmål*, *EF-strid*, *EF-tilhenger*. Jeg skal ikke kunne sige hvad der ligger til grund for at disse ord ikke er synlige i netordbogen, men i en historisk dokumentationsordbog er det uforståeligt hvis det er en overlagt beslutning. Visse af de førstnævnte ord tilhører sandt nok ikke sprogets mest centrale gloser, men i hvert fald *EF*-ordene er ikke sjældne i NAOB's periode. Måske – og forhåbentlig – er der en teknisk forkla-

ring på at ordene ikke har været tilgængelige i den periode jeg foretog mine opslag.

4 Præsentation af data: hjemmesiden

Hjemmesiden er lavet i et såkaldt responsivt design, hvilket vil sige at den automatisk tilpasser sig forskellige platforme og skærmstørrelser. Det er standard i dag, og jeg kan blot konstatere at det fungerer fint, også på små skærme som mobiltelefoner. Hvad der derimod ikke er standard og derfor fortjener særskilt omtale, er sidens hastighed. Det er ganske enkelt imponerende hvor hurtigt siden loader, og hastighed er et meget vigtigt parameter for succes hos brugerne. Det gør det til en fornøjelse at bruge ordbogen som dagligt redskab, og min vurdering er at dette alene vil gøre ordbogen til et standardværk for den der søger besked om norske ords betydning.

NAOB har et meget enkelt layout som gør det nemt at orientere sig på siden. Der er brugt få grafiske virkemidler og kun én central spalte. Af typografiske virkemidler er det bare skriftsnit (fed/ordinær/kursiv/kapitæler), farve (sort/rød, blå til link) samt i begrænset omfang skriftstørrelse der er taget i brug. Så vidt jeg kan bedømme, er der brugt samme font overalt (Arial). Brugen af rød kapitæltekst til alle metasproglige informationer er en modig og elegant løsning der stort set fungerer fint. Står den røde tekst i fed, fungerer den som overskrift for den følgende oplysningstype: bøjning, etymologi, betydning og brug, eksempler, citater og så videre. Står den i ordinær, bruges den i en række forskellige funktioner. Løsningen er dermed også en smule risikabel, idet den stiller krav til brugeren om korrekt afkodning af det flertydige budskab: Her vil redaktionen fortælle mig noget ekstra om opslagsordet i denne betydning. For at se på hvilke oplysninger de nærmere dækker over, kan vi se på nogle eksempler.

Tekst i rød kapitælskrift kan være en oplysning om overført betydning, om genre eller medium (ordsprog, mundtlig), om fag (tele, medicin, kemi osv.) og andre sprogbrugsmarkører (udover mundtlig også sjælden, dialektalt, historisk, litterært osv.). Andre gange gælder oplysningen en begrænsning i brugen: *skryte*, om esel: 'frembringe en høy, skrikende, hikstende lyd', *kvekke*, om frosk: 'frembringe en enstonig stadig gjentatt lyd', *klumpfotet*, om person: 'som har klumpfot'.

Det kan også være grammatiske oplysninger. NAOB stiller her ganske store krav til sine brugeres grammatiske kompetencer. Formuleringer som “med funksjon av determinativ” (udtrykket *i masse*), “som s-verb (med resiprok betydning)” (*kysse*) og “med oplever, resipient som indirekte objekt” (*fattes*) viser enten at de grammatiske kundskaber er på et væsentlig højere niveau i Norge end i Danmark, eller at NAOB’s kernebrugere er eksperter og højtuddannede.

Endelig bruges den røde tekst også til en særlig type forklaringer, nemlig den slags som ikke er direkte, udskiftelige definitioner af den traditionelle slags, men metadefinitioner. Metadefinitioner er en helt almindelig og legitim måde at definere på og bruges især ved den abstrakte del af ordforrådet, fx grammatiske funktionsord eller pragmatiske udtryk som hilsner, bandeord og kraftudtryk (jf. Lorentzen og Trap-Jensen 1989: 89 ff.). Et eksempel kan ses i figur 3.

fy faen | fy fanden

SOM UTTRYKK FOR SINNE, INDIGNASJON

- *en venn av familien [til Johan Borgen] hadde en gammel papegøye som kunne si en eneste setning, som gjengjeldt ble hyppig gjentatt: «Leser De Dagbladet, fy faen! Leser De Dagbladet, fy faen!» (Aftenposten 28.04.2002/Kultur/11)*
- *js. fy for fan' da!* (Henrik Ibsen: *En folkefiende* 142)

fakta faen

FORSTERKENDE NÅR MAN BEDYRER SANNHETEN AV NOE | jf. [faktum](#)

- *jeg har lest alle bøkene dine. Fakta faen. Alle sammen* (Sverre Henmo: *Dager uten sko* 164)

Figur 3: Metadefinitioner med rød kapitælskrift.

Den røde skrift dækker altså over en række forskellige typer af metasprog, men det er godt tænkt, og som bruger tror jeg godt man kan vænne sig til det. Til gengæld er det nødvendigt at være konsekvent, og her er endnu noget arbejde at gøre så man undgår eksempler på metadefinitioner med normal sort skrift sådan som eksemplet i figur 4 viser. Her kan være tale om en parsingfejl eller en menneskelig fejl, og det ville også være mærkeligt om alt var perfekt. Min oplevelse er at det ikke er typisk.

Som tidligere nævnt har redaktionen under omsætning af den typografisk opmærkede tekst til indholdselementer i XML-format forenklet NRO’s dybdehierarki så artiklerne fremstår mere overskuelige og tilgængelige. Den trykte ordbogs brug af romertal, bogstaver og flere niveauer af arabertal, foruden finere nuancer angivet ved hjælp af dobbeltstreger og semikolon, er afløst af en fladere struktur, dels gennem

gid fanden hadde ham

sagt om noe(n) man er rasende sint på (for)

- *gid fanden hadde 'n [futen] lys levende, den bondeflåeren!* (P.Chr. Asbjørnsen og Jørgen Moe: *Samlede eventyr 1*;

fanden ta ham

sagt om noe(n) man er svært ergerlig på (for)

- – *Faen ta deg. Hvorfor må du gjøre alt så vanskelig* (Levi Henriksen: *Snø vil falle over snø som har falt 172*)
- *foen ta USA! Slik kunne Harold resonnere mer og mer høylytt* (Lotta Elstad: *Et eget rom 151*)

Figur 4: Metadefinitioner bør ikke, som her, optræde i samme stil som normale definitioner.

den tekniske parsing af datastrukturen, dels gennem en manuel redaktionel tilpasning foretaget under revisionen af hver enkelt artikel.

Resultatet af forenklingen er i det store hele meget vellykket. Gevinsten er en langt større overskuelighed og en mere luftig og brugervenlig layout. Det er vel projektets største enkeltstående præstation at det er lykkedes så flot som det er tilfældet, og jeg spår at det – sammen med sidens hastighed – gør at ordbogen vil nå ud til en stor kreds af brugere. Her er der grund til at ønske redaktionen tillykke med en flot præstation.

5 Hjælp til søgning

Vores erfaring fra DDO er at brugerne ikke nødvendigvis slår op på lemmaformen når de vil finde et ord; det kan vi se både af vores søgelog og af de brugerundersøgelser vi har gennemført. I stedet søger brugerne efter den form de møder ordet i når de for eksempel læser en tekst. Det er derfor ofte bøjede former der søges efter, og her er det en god service hvis man får træf når man søger på en bøjet form. Jeg har prøvet forskellige muligheder for at se hvilken hjælp NAOB giver.

Hvis jeg søger på *bilen*, resulterer det i 6 træffere, heriblandt i hvert fald den ene eller de to jeg forventede, nemlig bestemt form ental af *bil* eller *bile*. Ud over de to forventede får man også træf på to homografartikler til *bil* (som jeg ikke kendte og derfor ikke forventede) samt verbet *bile* (vel fordi grundformen af substantivet *bile* falder sammen med verbet) og substantivet *bel* (vel fordi det er registreret som en variantform af *bil*). Jeg får altså træf på en hvilken som helst variant-, grund- eller bøjningsform der er registreret for et lemma der har den indtastede streng som én af sine former. Det gælder i øvrigt også former som ikke er eksplicit

nævnt i artiklen. En søgning på *bils* giver ligeledes træf på de relevante opslagsord, og *bilende* giver træf på verbet *bile*. Det er fint.

Derimod får jeg ikke noget træf hvis jeg søger på datidsformen *løp* af verbet *løpe*, og jeg antager at det er fordi der er et direkte træf på substantivet *løp*. En søgning på *smukke* bekræfter det: Det giver ikke træf på adjektivet *smukk* fordi der er træf på verbet *smukke*. Jeg vil anbefale at man også får træf på bøjningsformer af ord, også selvom der er et direkte træf. Man kan jo fint rangordne hittene således at direkte træffere vises før træffere på bøjningsformer.

En anden fin søgehjælp er ordfuldførelse, som aktiveres ved det andet indtastede bogstav, og der vises de første 10 nærmeste alfabetiske naboer. Det er både en hjælp til usikre stavere og ofte også tidsbesparende når man slipper for at indtaste hele ordet i søgefeltet. Det samme er søgehjælpen “Mente du”, som man kender fra Google, og som efterhånden er standard på mange sider. Den bagvedliggende algoritme kan dog ramme mere eller mindre hensigtsmæssigt, og også i NAOB er der forskel på hvor godt den fungerer. Så vidt jeg kan gennemskue, tæller den primært antal fælles bogstaver og foreslår opslagsord der har flest mulige bogstaver fælles med det indtastede. Derimod fejler den ofte ved en af de almindeligste indtastningsfejl, nemlig ombytning af to bogstaver. Søger jeg på *køjre*, får jeg foreslået *kipre*, *kikre*, *vejre*, *tømre*, *møje*, *kure*, *kurre*, *røre*, *kovre*, *kåre*, men ikke det (for et menneske) oplagte *kjøre*. Heller ikke den danske form *køre* giver det rigtige forslag, til trods for at formen uden *-j*- optræder i seks citater i artiklen.

I princippet er søgehjælpen altså fin, men algoritmen bør nok justeres lidt så den fanger flere af de oplagte former.

Endelig skal også nævnes muligheden for søgning i fritekst. Det er en udmærket funktion til søgning efter flere forekomster af ord eller udtryk eller efter andre oplysninger end opslagsord, f.eks. citater fra en bestemt kilde eller bestemte metasproglige bestemmelser (“med refleksiv betydning”, “poetisk”, “transitivt” osv.)

6 NAOB som historisk ordbog

NAOB dækker et ganske langt tidsrum, næsten 200 år, og det giver nogle særlige udfordringer for den bruger der søger information om det moderne sprog. Mens opslagsformer, bøjningsformer og redaktionelt

sprog bør være i overensstemmelse med gældende normering, giver det sig selv at det ikke er tilfældet for de mange citater der optræder i ordbogen. Det redaktionelle sprog følger Språkrådets normering for moderat bokmål fra 2005, og det er nødvendigt for at ordbogen kan bruges i skole- og undervisningssystemet for moderne norsk, men den normering gælder naturligvis ikke for de mange ældre litterære citater. Det skal brugeren være opmærksom på. En oplagt hjælp ville derfor være at oplyse årstallet for de anførte citater. Det gør NAOB for visse teksttyper, især aviser og internetcitater, men ikke for litterære citater, heller ikke de nyere. Hvis litteraturoversigten er samlet i databaseform, er det forholdsvis nemt at tilføje udgivelsesåret, og det vil jeg anbefale redaktionen at gøre.

7 Konklusion

Der er god grund til at lykønske redaktionen, Kunnskapsforlaget og Det Norske Akademi med NAOB. *Norsk Riksmålsordbok* er på flere måder et vanskeligt tilgængeligt værk. Værkets uegale fremtoning, fordelt på fire grundbind af stærkt varierende omfang og to supplementsbind, vidner om en skiftende redaktionel linje undervejs i tilblivelsen af det originale værk, og den komplicerede struktur og de valgte typografiske løsninger gør at værket ikke hører til de mest brugervenlige ordbøger at finde rundt i. Opgaven med at transformere det heterogene og vanskeligt tilgængelige forlæg til et moderne, brugervenligt opslagsværk er stort set lykkedes fint. Forandringen af det strukturelle dybdehierarki til en fladere og lettere tilgængelig struktur er vellykket, og de enkeltstående fejl der kan påvises, er i den forbindelse at regne for bagateller. Strukturændringen har gjort det muligt at skabe en enkel hjemmeside, som ikke alene imponerer ved sin overskuelighed og elegante typografi, men også med en hastighed og stabilitet der gør det til en fornøjelse at bruge NAOB, uanset hvilken platform det sker på.

At der er ting som kan forbedres, kommer næppe bag på redaktionen. Nogle af de ting jeg har fremhævet som mindre heldige, er de faste udtryks placering og ordning, algoritmen til søgehjælp og den manglende datering af litterære citater. Jeg er sikker på at redaktørerne kan supplere med en liste af ting de ønsker forbedret, og som de er klar til at tage sig af. Tilbage er kun at håbe at ordbogen tilføres de nødvendige driftsmidler

så den kan holde sig aktuell som samtidsordbog for bokmålet. For en netordbog som NAOB markerer lanceringen ikke projektets afslutning, men begyndelsen på en ny æra.

Litteratur

- Bojsen, Else og Lars Trap-Jensen. 2005. "ODS, ODS-S og fremtiden". I: *10. Møde om Udforskningen af Dansk Sprog*, red. Peter Widell og Mette Kunøe, 58–67. Aarhus: Aarhus Universitetsforlag.
- Bäckerud, Erik. 2014. "Nydigitalisering av SAOB". I: *Nordiske Studier i Leksikografi* 12. Skrifter utgitt av Nordisk forening for leksikografi, red. Ruth Vatvedt Fjeld og Marit Hovdenak, 95–105. Oslo: Novus Forlag.
- DDO = *Den Danske Ordbog*. <<https://ordnet.dk/ddo>>
- Lorentzen, Henrik og Lars Trap-Jensen (2012). "Nødvendigt, tilstrækkeligt, typisk? – nogle tanker om definitionspraksis". *LexicoNordica* 19: 79–101.
- NAOB = *Det Norske Akademis ordbok*. <<https://www.naob.no>>
- ODS = *Ordbog over det danske Sprog*. <<https://ordnet.dk/ods>>
- Nilstun, Carina. 2012. "Fra dype betydningshierarkier til flattere struktur – et innblikk i revisjonen av Norsk Riksmålsordbok". *LexicoNordica* 19: 123–140.
- Svendsen, Mette Marie Møller. 2016. *Brugerundersøgelserapport for Den Danske Ordbog*, intern rapport. København: Det Danske Sprog- og Litteraturselskab.

Lars Trap-Jensen
 Det Danske Sprog- og Litteraturselskab
 Christians Brygge 1
 DK-1219 København K
 ltj@dsl.dk