

Särtryck ur

Nordiska studier i lexikografi

11

Rapport från
Konferensen om lexikografi i Norden
Lund 24–27 maj 2011

Netordbøger – hvordan finder
brugerne frem til dem, og hvad gør
de, når de har fundet dem?

Henrik Lorentzen & Liisa Theilgaard

Skrifter utgivna av Nordiska föreningen för lexikografi
Skrift nr 12

Netordbøger – hvordan finder brugerne frem til dem, og hvad gør de, når de har fundet dem?

Henrik Lorentzen & Liisa Theilgaard

In general, user behaviour studies of online dictionaries have focused on user behaviour once the user is on the site. But before a potential user even reaches this stage, he or she must succeed in finding the dictionary on the Web. The purpose of this paper is to investigate users' linguistic search strategies before they enter our dictionary site, ordnet.dk. What kind of search engine queries are successful and why (not)? Similarly, we have studied the queries in the search field of our online dictionaries. Are the search strategies the same as when using a search engine? Taking the no-match searches as a starting point, we have asked ourselves whether our content and search functionality correspond to the search behaviour of the users, i.e., if we can give an answer to the users' query and if data are organized and presented in an appropriate way. Given the results of these analyses, we decided to make several changes to the site in order to optimize user access and attract new users. We present these changes and the results and finally some ideas about future development. The paper is based on empirical data from Google Analytics and our log files.

Nøgleord: netordbøger, brugeradfærd, søgestrategier, webstatistik, søgemaskineoptimering.

De fleste undersøgelser af brugeradfærd i forbindelse med netordbøger har koncentreret sig om adfærden, når brugeren er kommet ind på ordbogssitet. Men inden en potentiel bruger overhovedet er nået dertil, må han eller hun finde frem til ordbogen på nettet.

I denne artikel undersøger vi de søgestrategier, brugerne anvender, når de leder efter sproglig viden ved hjælp af en søgemaskine. Med resultaterne af disse undersøgelser i hånden har vi stillet os selv spørgsmålet: Hvordan kan vi optimere Det Danske Sprog- og Litteraturselskabs netordbøger på ordnet.dk, så de tilgodeser denne søgeadfærd bedst muligt og hjælper flest mulige brugere frem

til ordbogen og dermed til et svar? Vi præsenterer nogle af de ændringer, vi har foretaget, og svarer på, om de har givet resultat i form af flere nye brugere, flere opslag med videre.

På tilsvarende vis har vi undersøgt, hvad brugeren søger på i netordbøgernes søgefelt. Er søgestrategierne de samme som i søgemaskinerne, og er brugeren overhovedet bevidst om, hvad Den Danske Ordbog og Ordbog over det danske Sprog er for en slags ordbøger? Med udgangspunkt i de søgninger, der ikke har givet resultat, har vi spurgt os selv, om vores indhold og søgefunktion stemmer overens med brugernes søgeadfærd, dvs. om vi kan svare på brugerens søgning, og om data er organiseret og præsenteret hensigtsmæssigt. Det giver desuden anledning til at overveje, hvilke veje vi skal gå i fremtiden for at supplere netordbøgerne med nyt og andet indhold.

Metode, termer og redskaber

Webstedet ordnet.dk hviler i øjeblikket på tre søjler: Den Danske Ordbog (DDO), Ordbog over det danske Sprog (ODS) og KorpusDK (se Fjeld og Malmgren 2010). Den første netversion af ODS blev lanceret i november 2005, og derefter fulgte DDO i november 2009. I november 2010 lanceredes anden version af ODS, hvor hele, nye artikler fra supplementet er inkluderet. KorpusDK muliggør søgninger i to korpuser over moderne dansk almentsprog og er til en vis grad integreret med ordbogsressurserne, men vil i øvrigt ikke blive berørt i denne artikel.

I undersøgelserne af brugernes adfærd har vi primært benyttet to redskaber, dels Google Analytics, dels vores egne logfiler. Google Analytics er et online webanalyseværktøj (se Clifton 2010), der leverer en lang række statistikker over de besøgendes adfærd. Disse statistikker oplyser fx om, hvor mange besøgende siden har pr. dag, og om, hvordan de besøgende er kommet frem til siden. Det er muligt at sammenligne forskellige perioder med hinanden, fx den samme måned i to forskellige år, eller se, hvor længe en bruger opholder sig på siden.

Samtidig holder vi også selv øje med brugerne i kraft af vores egne logfiler (jf. bl.a. Almind 2008 og Nygaard og Fjeld 2008), nærmere betegnet via de såkaldte no match-lister, dvs. lister over, hvad brugeren *ikke* finder ved sine søgninger på ordnet.dk. De undersøgte data gælder DDO, dels fordi det er den ordbøgerne, der har langt flest besøgende, dels fordi det er den ordbogsbase, vi primært opdaterer, men det er muligt at foretage de samme undersøgelser for ODS.

I undersøgelsen opererer vi med et “før” og “efter”. “Før” er september 2010, og “efter” er marts 2011. I den mellemliggende periode har vi foretaget

en række ændringer i webstedets funktionalitet, og det er derfor muligt at se, om disse ændringer har haft nogen indvirkning på besøgstal og brugeradfærd.

Når vi taler om “besøg”, er der tale om antallet af individuelle sessioner, som en bruger har startet. Hvis brugeren har været inaktiv i 30 minutter eller derover, regnes ny aktivitet som et nyt besøg.

Hvordan finder brugerne frem til ordbogen?

Der er tre forskellige adgangsveje til ordbogen:

- 1) Brugeren kan søge i en søgemaskine og klikke på et link til ordbogen i søgeresultatet
- 2) Brugeren kan komme direkte til ordbogen ved at klikke på et bogmærke gemt i browseren eller ved at taste webadressen direkte i browserens adresselinje
- 3) Brugeren kan komme via en linkhenvisning på et andet websted

I september 2010 kom halvdelen (49 %) til DDO ved at søge i en søgemaskine – og heraf står Google for 99 % – en tredjedel (33 %) kom direkte til DDO, og de sidste 18 % kom via en henvisning på et andet websted, fx sproget.dk. Henvisningsbrugere bruger kortest tid pr. besøg (5:31 min.) og ser færrest sider pr. besøg (3,53 sider). De flittigste kommer direkte og bruger 7:21 min. og ser 4,1 side pr. besøg, og det undrer ikke, da de kender webstedet og ordbogen i forvejen.

Brugerne foretrækker altså at søge i en søgemaskine, og dermed bliver det meget vigtigt at figurere som søgeresultat i Google – og helst på den første søgeresultatside – hvis en potentiel bruger skal blive opmærksom på ordbogen.

Hvordan søger brugerne i en søgemaskine?

For at undersøge brugernes søgeadfærd i en søgemaskine har vi undersøgt alle de søgeudtryk, som brugerne har søgt på i en søgemaskine i september 2010, og som har resulteret i et klik videre til ordbogen. Undersøgelsen siger altså *ikke* noget om, hvilke søgeudtryk der mislykkedes, dvs. ikke ledte frem til ordbogen, eller hvilke søgeudtryk brugeren eventuelt har forsøgt sig med inden.

Først har vi set på antallet af søgeord i søgeudtrykket, defineret som en streng uden mellemrum. Her viser det sig, at 60 % af brugerne har skrevet ét søgeord, 23 % to søgeord og 17 % tre eller flere ord, dvs. at hele 40 % søger på mere end ét ord.

Dernæst har vi grupperet alle søgeudtryk i tre kategorier, alt efter om det betegner et emne (metaord, fx *ordnet*, *dansk ordbog*), et sprogligt udtryk (materialiterord, fx *slesk*, *få blod på tanden*) eller en kombination af disse to (fx *hvilken ordklasse er ordet efter* eller *nonchalant betydning*, hvor brugeren søger efter betydningen af ordet *nonchalant*):

	Antal søgeudtryk	Antal besøg	Besøg i %	Eksempler
Metaord	1.160	28.355	64 %	<i>ordnet</i> , <i>dansk ordbog</i> , <i>ddo</i>
Materialiter	2.220	13.290	30 %	<i>slesk</i> , <i>selvom selv om</i>
Kombination	1.251	2.780	6 %	<i>nonchalant betydning</i> , <i>snyde synonym</i>
I alt	4.631	44.425	100 %	

Figur 1: Fordeling af søgeudtryk via en søgemaskine i september 2010

Søgeudtrykkene fordeler sig med 2.220 materialitersøgeudtryk, 1.160 metasøgeudtryk og 1.251 søgeudtryk, hvor metaord er kombineret med materialiterord. Disse 4.631 forskellige søgeudtryk har resulteret i 44.425 besøg, dvs. i gennemsnit 10 besøg pr. søgeudtryk. Men der er stor forskel på effektiviteten af søgeudtrykkene. Således har 2.530 søgeudtryk kun genereret ét besøg hver, svarende til 5,7 % af besøgene. Heroverfor står de godt 1.160 metaord, der har genereret 64 % af besøgene, hvoraf de fem mest effektive søgeudtryk *ordbog*, *dansk ordbog*, *den danske ordbog*, *ordnet* og *ordnet.dk* tegner sig for hele 39 % af alle besøg. Det er altså i høj grad søgeudtryk med metaord, der har givet resultat, og det hænger sammen med, hvad der er blevet indekseret af webcrawlerne.

Ved lanceringen af *ordnet.dk* blev det af frygt for manglende serverkapacitet besluttet udelukkende at tillade indeksering af forside, hjælpesider og omtekster, dvs. omkring 25 sider fra DDO og omkring 65 sider i alt på *ordnet.dk*. Der var således ingen ordbogssider, der blev indekseret automatisk, men enkelte ordbogssider er alligevel blevet indekseret, fordi andre websteder har linket til en bestemt ordbogsside. Denne praksis harmonerer åbenbart dårligt med brugernes faktiske søgeadfærd, så der ligger et stort potentiale i at kunne tilgodese materialitersøgning fra en søgemaskine.

Hvordan søger brugeren i søgefeltet?

Sammenligner vi brugernes søgeadfærd i en søgemaskine med søgeadfærden, når brugeren er kommet ind på DDO og søger i søgefeltet, tegner der sig et ganske andet billede.

Til denne delundersøgelse har vi måttet bruge data fra marts 2011, fordi logningen af søgeudtryk fra søgefeltet ikke var slået til i Google Analytics i september 2010. Her viser det sig, at 96 % søger på ét ord, mens 2,4 % søger på to ord og 1,6 % på tre eller flere ord. Set i forhold til de 60 % étordssøgninger i en søgemaskine er det tydeligt, at brugeren har en anden adfærd, når det gælder søgninger i søgefeltet.

Eftersom søgninger via søgefeltet i marts 2011 udgjorde så meget som 641.140 forskellige søgeudtryk, har vi ikke kunnet behandle dem kvantitativt ved en systematisk klassifikation, men et blik ned gennem søgeudtrykkene viser, at de fleste søger på et materialiterord. Søgning med jokertegn er sjældent brugt, idet kun 0,3 % af søgeudtrykkene indeholder et jokertegn (fx **skib, ???tøj*). Vi har ellers gjort meget ud af, at man kan søge på både udtryk og flerordsforbindelser samt med jokertegn, men tilsyneladende har brugerne ikke lagt mærke til det eller følt det som et behov.

Meta- og kombinationssøgninger forekommer heller ikke særlig ofte, fx optræder strengen *ordbog* kun 144 gange, *synonym* 56 gange og *betydel betydning* 46. Brugeren opfatter altså søgefeltet som et opslagsfelt til ordbogsdelen alene, og sådan fungerer søgefeltet også normalt i elektroniske ordbøger i modsætning til andre websteder, hvor det bruges til fritekstsøgning på hele hjemmesiden.

Hvad finder brugeren ikke?

Ved hjælp af vores logfiler har vi undersøgt samtlige registrerede søgninger via DDO's søgefelt i september 2010, og heraf var der 20,6 %, der ikke gav noget resultat. Dette tal er lavt sammenlignet med to tidligere undersøgelser, hvor ikke-vellykkede søgninger i Bokmålsordboken udgjorde 31 % (Nygaard og Fjeld 2008: 59) og i Lexin online Svenska ord udgjorde 41 % (Hult 2008: 78), hvilket i det sidste tilfælde måske hænger sammen med, at ordbogen indeholder så få opslagsord, og at ordbogen har brugere med svensk som andetsprog som hovedmålgruppe.

Man skal være opmærksom på, at dette tal kun dækker søgninger uden direkte resultat i form af en (ordbogs)side, og at det derfor ikke siger noget om, hvor mange af disse søgninger der resulterede i et meningsfuldt mente du-for-slag. Vi har nemlig truffet det valg, at brugeren kun ledes direkte til en ordbogs-side, når der er tale om bøjningsformer og ortografiske varianter, der er registre-

ret i ordbogsartiklerne. Ellers præsenteres brugeren for en række mente du-forslag, der dels bygger på den søgealgoritme, vi anvender¹, dels på en liste med kendte fejlformer.

Herefter har vi kategoriseret de 1.000 mest frekvente ikke-vellykkede søgeudtryk i en række typer. Disse 1.000 søgeudtryk repræsenterer 12,6 % af samtlige ikke-vellykkede søgeudtryk. Et søgeudtryk kan optræde i flere kategorier, fx både som en stavfejl og som en bøjningsform. Den mest almindelige ikke-vellykkede søgetype var stave- og slåfejl med 30 % (fx *unskyld*, *priotere*, *sjangre* for *genre*), og dernæst kom bøjningsformer med 20 % (fx *dominerende*, *terroristernes*, *bibringes*, *skrid*, *fornemste*). Det overraskede os, at der blev søgt på så mange bøjningsformer. En mulig forklaring er, at vi tilbyder en opensearch-funktion. Med den installeret i sin browser (Explorer eller Firefox) kan man bruge DDO som søgetjeneste og slå et ord op direkte i ordbogen. Man markerer et ord på en hvilken som helst webside og vælger "Søg med Den Danske Ordbog", og så sendes det markerede ord af sted som søgning til ordbogen og vises på et nyt faneblad. Vi har dog ikke nogen tal for, hvor mange der benytter denne funktion, så vi ved det ikke. Under alle omstændigheder fik det os til at tænke over, hvilke bøjningsformer vi ikke havde taget højde for i vores data.

Resten fordeler sig især på typer som fremmedord (9 %), proprier (7 %), sammensætninger (7 %), nye ord (6 %), flerordsudtryk (4 %), forkortelser (3 %), afledninger (2 %) samt flere andre, og de siger selvfølgelig noget om, hvad brugerne leder forgæves efter i DDO (se mere nedenfor om fremtidsperspektiver).

Ændringer

På baggrund af vores undersøgelser besluttede vi at foretage en række ændringer. Først og fremmest ønskede vi større synlighed i søgemaskinernes søgeresultater – og helst på første resultatside – og derfor åbnede vi for, at webcrawlere kunne indekserer alle de enkelte ordbogssider. Så hvor der før var under 100 indekserede sider, er der nu ca. 100.000 sider for DDO og 350.000 for hele ordnet.dk.

Dernæst skulle der gøres noget ved de mange bøjningsformer, som vi ikke var klar over, at der blev søgt på. Vi har derfor tilføjet præsens participium som skjult bøjningsform til verbernes bøjning. Der er sikkert mange, der søger på

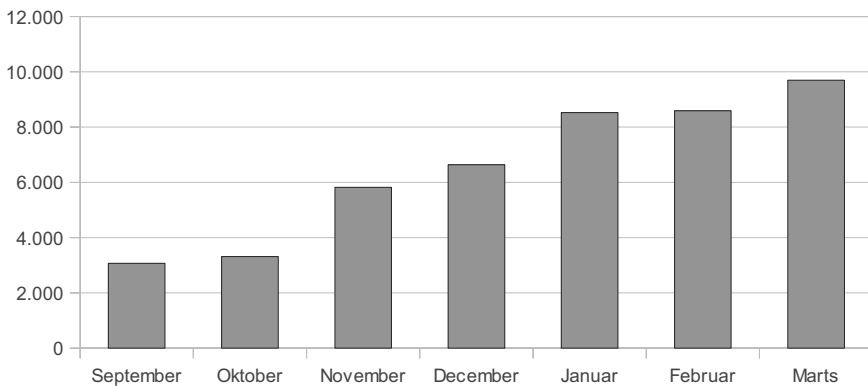
¹ Det drejer sig om stavekontrolprogrammet *Aspell* i kombination med *Levenshtein*-distancen, som er et mål for forskellen mellem to strenge, og som vi bruger til at sortere de dårligste forslag fra.

denne bøjningsform, fordi de bruges som adjektiver, men i DDO beskrives de ofte i verbalartiklen og ikke i en selvstændig artikel. Så nu kommer man automatisk til artiklen *dominere*, når man søger på *dominerende*. På samme måde har vi genereret skjulte genitivformer til alle substantiver, så man nu automatisk kommer til artiklen *trommehinde*, når man søger på *trommehindens*. – Til gengæld har vi ikke gjort noget for at tilføje imperativ og passiv, da disse former ofte genkendes af søgealgoritmen bag mente du-funktionen. Adjektivernes komparationsformer er medtaget som bøjningsformer i DDO, men bestemt form superlativ (*den/det/de fornemste*) er ikke medtaget, og den har vi heller ikke gjort noget ved endnu.

Endelig har vi gjort mente du-funktionen mere synlig ved også at vise den i midterfeltet og ikke blot under søgefeltet, så der kom mere fokus på de ordforslag, der skal hjælpe brugeren, når en søgning ikke giver resultat. Desuden har vi forbedret mente du-funktionen, så den håndterer sammenskrevne former bedre og viser et henvisningslink, når ens søgeudtryk findes i ODS, men ikke i DDO.

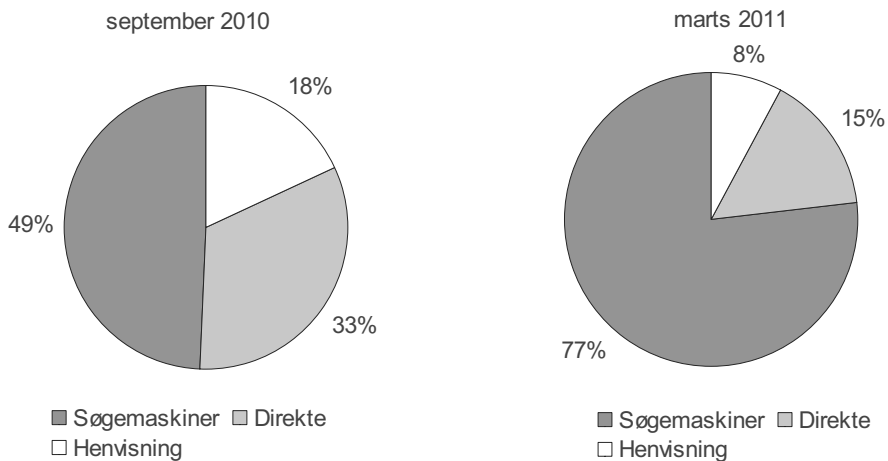
Resultater

Indekseringen af alle ordbogssider har haft en markant effekt. Antallet af daglige besøg er mere end tredoblet i løbet af det halve år, vi har undersøgt trafikken, og det fortsætter.



Figur 2: Besøg på DDO pr. dag i perioden september 2010 til marts 2011

Dette afspejler sig naturligvis også i fordelingen af brugernes adgangsvej til ordbogen. Hvor det i september 2010 var 49 % besøg, der kom fra en søgemaskine, er det i marts 2011 hele 77 %:



Figur 3: Adgangsveje til DDO i september 2010 og marts 2011

Indekseringen har altså betydet, at vi nu figurerer dér, hvor brugerne søger efter information (også sproglig), nemlig i søgemaskinerne (Google) – og ofte på første søgeresultatside. Det gælder både de nye og de tilbagevendende brugere, idet 82,3 % af de nye brugere i marts 2011 kommer via en søgemaskine, men også 72,6 % af de tilbagevendende. Samtidig bidrager søgemaskinernes egne søgealgoritmer til at guide brugerne frem til et resultat, for eksempel i forhold til stavfejl, bøjningsformer og lignende. Dertil kommer, at brugerne nu henvises direkte til svaret i form af en ordbogsside og ikke bare til forsiden. Og endelig bør det vel nævnes, at denne synlighed ikke har kostet en krone til webmarkedsføring eller lignende, og at webcrawlnes indeksering ikke har betydet længere svartider.

Andelen af nye brugere er også steget, så fordelingen mellem nye og tilbagevendende besøgende i marts 2011 er 44 % nye over for 56 % tilbagevendende, mens det i september 2010 var 34 % nye over for 66 %. Det er dog stadig mere end halvdelen, der er tilbagevendende besøgende, og de bruger stadig længst tid på hvert besøg. Opgørelserne viser, at det altså både er lykkedes at tiltrække nye brugere og at gøre en del af dem til trofaste brugere.

Ser vi på de søgeudtryk, som har ledt brugerne frem til ordbogen via en søgemaskine, er der nu fem gange så mange, der kommer til ordbogen via et søgeresultat, men samtidig er der blevet 18-19 gange så mange forskellige vellykkede søgeudtryk, nemlig 85.601 i forhold til septembers 4.631, dvs. at der er en langt større spredning, når det gælder søgeudtryk. Det hænger sammen med, at alle ordbogssiderne nu bliver indekseret, så der er chance for at få resultat med et langt større antal søgeudtryk. Det kan illustreres ved, at der i september

2010 var 2.530 forskellige søgeudtryk, der resulterede i hver 1 besøg, svarende til 55 % af alle søgeudtryk, mens det i marts 2011 var 60.384 svarende til 70,5 %. Noget tilsvarende ses i fordelingen på søgeudtrykstyper: Hvor det før i høj grad var metasøgeudtryk som *ordbog*, der gav resultat, er der nu meget større sandsynlighed for at få match med et materialitersøgeudtryk, fordi opslagsordet indgår i navnet på de enkelte websider med ordbogsartikler. Alt i alt har webcrawlernes indeksering af ordbogssiderne været den vigtigste enkeltfaktor, når det gælder om at tilgode se brugernes søgeadfærd: De søger sproglig information via en søgemaskine, og de søger ofte med materialitersøgeudtryk.

Tilføjelsen af yderligere bøjningsformer har også givet resultat. Af de 1.000 mest frekvente ikke-vellykkede søgninger fra september 2010 kan alle bøjningsformer undtagen passivformen *fåes* og imperativformen *ryd* håndteres, enten via de tilføjede bøjningsformer eller via mente du-funktionen, idet søgealgoritmen kan foreslå det rigtige ord, når det gælder mange af stave- og slåfejlene. Samtidig kan vi henvise til ODS, hvor der ikke er resultat i DDO. ODS har 225.000 opslagsord og dækker en ældre periode, så der henvises for eksempel, når der er tale om ældre ord som *møgle*, sammensætninger som *fødselsoverskud* og fremmedord som *sekularisme*. Det betyder, at af alle de 1.000 mest frekvente ikke-vellykkede søgninger fra september 2010 kan vi nu håndtere de 45 %, hvoraf de 10 % er henvisninger til ODS. Hvis vi antager, at resultatet for de frekvente kan overføres til også at gælde alle ikke-vellykkede søgninger, kan vi altså håndtere 45 % eller næsten halvdelen af de i alt 20 % ikke-vellykkede søgninger, dvs. at kun ca. 10 % af samtlige søgninger ikke kan håndteres, og det synes vi er et meget tilfredsstillende resultat.

Fremtidsperspektiver

Undersøgelserne af logfilerne og statistikkerne fra Google Analytics giver os mange idéer til fremtidigt arbejde, dels med hensyn til indhold, dels med hensyn til funktionalitet: Hvad kan og skal brugeren gøre på siderne?

Fremmedord er en efterspurgt kategori, jf. følgende liste over ord, brugerne ikke kan slå op: *denotativ, emergent, residual, ebrieret, dopamin, intentionalitet, eksplorativ, inkrementel, semiotisk, nominal, elaborere*. Her kan redaktionen med fordel supplere. DDO er lidt underforsynet, hvad angår lidt sjældnere fremmedord på grund af de relativt strenge krav om korpusbaseret, som redaktionen har fulgt. Mange ord af denne type var ikke særlig frekvente eller overhovedet med i DDO's korpus. Dertil kommer, at mange brugere næppe slår udbredte dagligdags ord op, men mere ord, de ikke kender. Det skal dog tilfø-

jes, at man i DDO godt kan finde forklaring på flere af de nævnte ord via et beslægtet opslagsord, fx *semiotik* og *intentionel*.

En anden efterspurgt type er proprier. De er principielt slet ikke med i DDO, undtagen hvis de indgår i faste udtryk som *fra Herodes til Pilatus*. Der kan være flere grunde til, at proprier slås op: Der kan være problemer med stavningen: *Ålborg*, *Taastrup* (å eller aa), *Svejts*, *Egypten* (kan også staves *Schweiz* og *Ægypten*), eller der kan være tale om et vidensbehov, som vi kun kan gætte på: *Orion* (stjernebilledet?), *Luna* (?), *Nato* (hvad er det, og kan det staves *NATO*?), *Anders* (hvad betyder navnet?).

Sammensætninger er ikke overraskende også almindelige på listen over ikke-vellykkede søgninger: *bilreparation*, *bryllupsår*, *juletræstøppe*, *sigøjnerbande*, *koncessionsaftale*, *lømmelpakke*, *svenskerfrisyre*, *københavnskort*. Dansk, svensk og norsk er sprog, der kan danne i princippet uendeligt mange sammensætninger, og mange brugere vil måske bare slå et ord op for at se, om det findes, dvs. er registreret. Ord som de nævnte afspejler, hvad der tales om i tiden, fx *bryllupsår* og *sigøjnerbande*, og er som sådan ad hoc-dannelser, mens andre er udtryk for leksikalisering, fx *juletræstøppe*, der betegner en helt bestemt genstand. I *lømmelpakke* er ordet *pakke* leksikaliseret i en nyere betydning om et politisk tiltag, her et, der giver politiet udvidede beføjelser over for potentielt kriminelle. *Svenskerfrisyre*, der åbenbart indtastes sådan (korrekt stavning er *frisure*), er sammen med *svenskerhår* etableret på dansk som betegnelse for en bestemt frisure til mænd (kort hår fortil og langt bagtil; kaldes også *bundesligahår*).

Listerne byder også på helt nye ord, dvs. ord fra de sidste 10-15 år, som fx *slush ice*, *vintage*, *avatar*, *kvindechauvinisme*, *wasabi*, *zumba*, *simkort*, og af disse er de tre sidste da også kommet med i DDO inden for det sidste halve år.

Forkortelser kan man også godt forstå, at brugere vil slå op, og her kan der også suppleres med nyt stof: *LOL* (laughing out loud), *ifm* (i forbindelse med), *CI* (cochlear implantat), *ATP* (Arbejdsmarkedets Tillægspension). DDO indeholder allerede en del forkortelser, men området er naturligvis ikke udtømt, og der kommer nyt til, fx fra sms- og chatsproget. Mange forkortelser er proprier som fx *FN*, *EU*, *SNG*, *SS*, og også her kan redaktionen vælge at medtage denne type fremover, ikke mindst fordi de ofte er flertydige som *SS*: *Schutzstaffel*, *dampskib* (steamship) eller chatforkortelse for "smiler sødt".

Når det gælder synonymy, må det siges, at DDO opererer med et snævert synonymbegreb, dvs. at den generelt ikke indeholder så mange synonymy som fx *Nudansk Ordbog* og slet ikke så mange som en regulær synonymordbog, men brugernes indtastninger tyder på, at de ønsker flere synonymy. På *DSL* arbejdes i disse år med en ny begrebsordbog for dansk, baseret på DDO og på det danske wordnet *DanNet*, og det er vores plan engang i fremtiden at kunne

tilbyde flere synonymer og beslægtede ord på grundlag af data fra begrebsordbogen (se Nimb 2011 og Trap-Jensen 2011).

Det er også vigtigt og oplagt, at redaktionen arbejder videre med at forbedre søgning på ét eller flere ord og på mente du-funktionen, der hjælper brugere, som begår mindre taste- eller stavefejl. I øjeblikket får man ikke hjælp, hvis man indtaster en bøjningsform af en almindelig stavefejl som fx *paranteser*, mens man bliver ledt på rette vej ved indtastning af singularisformen *parantes* (for *parentes*). Her er det oplagt at udvide bøjningsleksikonet med bøjningsformer til stavefejl. Andre indtastninger ligger for langt fra det rigtige til, at de kan opfanges af mente du-funktionen, selvom stavemåden *angsement* afspejler normal dansk udtale af *arrangement*, og her vil vi gerne på længere sigt give mulighed for fonetisk søgning, sådan som det allerede længe har været kendt i flere udenlandske ordbøger (fx den franske TLFi). Vi kan også overveje at tilføje bøjningsformer som passiv og imperativ; no match-listerne indeholder som nævnt mange bøjningsformer, hvilket tyder på, at brugerne enten kopierer en ordform ind i søgefeltet eller bruger de muligheder for direkte opslag, som vi tilbyder (den tidligere omtalte opensearch-funktion).

Andre fremtidsperspektiver går på at udvide og forbedre hjemmesidens funktionalitet. I øjeblikket er idiommer og andre faste udtryk søgbare, så de kommer frem i søgeresultatet, hvis man indtaster ét eller flere af de ord, der indgår i udtrykket. Det gælder dog kun de faste udtryk, der internt kaldes sublemmaer; de står i en særlig sektion af artiklen og er forsynet med mange af de samme oplysninger som hovedopslagsordene. DDO indeholder imidlertid også et stort antal ukommenterede ordforbindelser (kollokationer i bred forstand), der blot er nævnt under de relevante betydninger. De findes ikke frem med de nuværende søgemuligheder, og det vil være en oplagt forbedring, hvis de kommer til det. Brugere kan nemlig ikke vide, om en given ordforbindelse står som fast udtryk eller kollokation – og kender nok heller ikke forskel. I øjeblikket finder man fx ikke følgende ordforbindelser: *i forbindelse med*, *heller ikke*, *akademisk uddannelse*, *familiens sorte får*, *gøre honnør*, *veje for meget*, og det er da ærgerligt, når de faktisk er taget med i ordbogen.

Endelig er der perspektiver i at se mere på brugernes metasøgeudtryk, fx *synonym*, *andet ord for*. Det tyder – som nævnt – på, at brugerne ofte er interesserede i synonymer, og hvis de er det, bør vi give synonymer en mere fremtrædende plads i præsentationen og også gerne tilbyde flere synonymer. Et søgeudtryk som *ord med p* kunne antyde, at hjælpeteksterne bør være bedre til at lære de besøgende at bruge jokertegn, eller at vi skal kunne håndtere natursproglig søgning i et eller andet omfang, fx så man kan stille et spørgsmål.

Konklusion

Vi kan således drage flere positive konklusioner på vores foreløbige erfaringer med brugerundersøgelser: For os har det været en succes at tillade webcrawlerne at indeksere alle vores sider, så vi kan kun opfordre andre til at prøve det. Besøgstallene på ordnet.dk er mere end tredoblet, siden vi begyndte i november 2010, og de er stadig stigende pr. september 2011, uden at vi har oplevet længere svartider eller nedbrud. Andre har før os konstateret, at logfiler giver værdifuldt input, og det er meget vigtigt at få dette både redaktionelle og funktionelle input. Helt sikkert er det i hvert fald, at der er meget arbejde at gøre fremover, og en væsentlig del af opgaven bliver at prioritere indsatsen.

KILDER OG LITTERATUR

- Almind, Richard, 2008: Søgemønstre i logfiler. I: *LexicoNordica* 15. S. 33–55.
- Becker-Christensen, Christian m.fl., 2010: *Politikens Nudansk Ordbog*, 21. udg. København.
- Clifton, Brian, 2010: *Advanced Web Metrics with Google Analytics*, 2. udg., Wiley Publishing, Inc. Indianapolis.
- DanNet = DanNet 2009-. København: Det Danske Sprog- og Litteraturselskab og Center for Sprogteknologi, Københavns Universitet: www.wordnet.dk (oktober 2011)
- DDO = Hjorth, Ebba, Kjeld Kristensen m.fl.: *Den Danske Ordbog 2003-*. København: Det Danske Sprog- og Litteraturselskab. <http://ordnet.dk/ddo> (oktober 2011)
- Fjeld, Ruth Vatvedt og Sven-Göran Malmgren, 2010: Vård ett besök – om DSL:s nya webbsida ordnet.dk. I: *LexicoNordica* 17. S. 297–311.
- Hult, Ann-Kristin, 2008: Användarna bakom loggfilerna – redovisning av en webbenkät i Lexin online Svenska ord. I: *LexicoNordica* 15. S. 73–91.
- Nimb, Sanni, 2011: Semantiske relationer i en ny dansk begrebsordbog: genbrug på tværs af ordbøger. I: *LexicoNordica* 18. S. 135–155.
- Nygaard, Lars og Ruth Vatvedt Fjeld, 2008: Analyse av søkelogger for bedre søkmu-ligheter i elektroniske ordbøger. I: *LexicoNordica* 15. S. 57–72.
- ODS = Dahlerup, Verner m.fl.: *Ordbog over det danske Sprog inkl. supplement 1918-2005*. København: Det Danske Sprog- og Litteraturselskab. <http://ordnet.dk/ods> (oktober 2011)
- sproget.dk: <http://sproget.dk> (oktober 2011)
- TLFi = Le Trésor de la Langue Française informatisé. <http://zeus.atilf.fr> (oktober 2011)
- Trap-Jensen, Lars, 2011: Dansk begrebsordbog: baggrund og metode. I: *LexicoNordica* 18. S. 183–197.

Henrik Lorentzen

Det Danske Sprog- og Litteraturselskab
hl@dsl.dk

Liisa Theilgaard

Det Danske Sprog- og Litteraturselskab
lt@dsl.dk