

*Bernt Møller**

Internet som tekstsamling og ressourcebank i sprogundervisningen

Abstract

The World Wide Web contains huge amounts of documents in several languages. Also, internet offers numerous resources which may help in finding and understanding web documents. This article firstly aims at passing on facts and considerations from the author's teaching experience within a thematic process framework of Information Management. Here, information searching, evaluation, registration and presentation are guidelines. In information searching, the points touched upon include search engine strategies, such as qualified URL guessing and language checking. Information evaluation mentions a few basic techniques while information registration and presentation point to the utility of co-integrating multimedia links in the software applications. Secondly, the article in a smaller section suggests a typology for internet language resources and comments on a number of useful language web sites. Finally, selected perspectives and challenges are outlined.

1. Indledning

"IT og sprog – få øje på kombinationen" lyder titlen på en kronik i Computerworld (Arleth 2001), der har til hensigt at gøre opmærksom på, at hvis IT skal være for alle, skal den tilhørende dokumentation også være skrevet i et sprog, der er tilgængeligt for alle. Belyst ud fra cand.ling.merc. Nina Arleth's erfaringer med manualsproget er dette som regel langt fra tilfældet. Som underviser i teknisk/teknologisk sprog og oversættelse på BA erhvervsprog og cand.ling.merc.-studiet samt i tværproglige moduler i IT-værktøjer under Åben Uddannelse og i IT-Vest-regi på Handelshøjskolen i Århus kan jeg – som sikkert mange pc-udstyrede og dermed manualbetyngede brugere – ikke genkende.

* *Bernt Møller*
Handelshøjskolen i Århus
Fransk Institut
Fuglesangs Allé 4
DK-8210 Århus V
bcm@asb.dk

IT-sproget kan der skrives meget om. Denne artikel har imidlertid ikke til hensigt at analysere IT-genrer eller IT-terminologi, men angriber *IT og sprog* ud fra en anden synsvinkel. Jeg ønsker at sætte fokus på internet belyst ud fra mine erfaringer siden 1996–97 med inddragelse af dette medie i sprogundervisningen. Artiklen er ikke tænkt som en systematisk gennemgang af internetmediets faciliteter, ej heller som en oversigt over gode webadresser eller over sprogdata-baser. Hensigten er derimod at præsentere og diskutere strategier og ressourcer, hvor såvel jeg som underviser og mine studerende har haft gavn af nettet op til i dag. Dette til mulig inspiration for sprogundervisere og sprogstuderende, der ønsker en måske lidt mere systematisk og reflekteret brug af internet i deres undervisning.

Siden 1996–97 har internetteknologien gennemgået ændringer, ikke mindst afspejlet i nettets stigende brugervenlighed. Tilsvarende – men endnu hurtigere – sker der markante ændringer i udbud og kvalitet af de tekster og andre ressourcer, nettet tilbyder. Set i lyset af det omfang, jeg og mine fortløbende studerende benytter det nye medie, samt af deres og min erfaringsopbygning, vil en kronologisk systematik i denne artikel være uhensigtsmæssig. Den vil i stedet blive tematisk og processuel. Artiklen tager i afsnit 2 og 3 afsæt i de studerendes behov og i processer inden for Information Management (IM), hvorefter fokus er på internetværktøjer til informationssøgning og -vurdering, især søgemaskinerne (afsnit 4 og 5). Afsnit 6 og 7 behandler systematiseringen af den udvalgte information i termbaser samt IT-støttet informationspræsentation. I afsnit 8 præsenteres udvalgte typer internetressourcer, såsom sprogportalen, opslagsværker og programmer til onlineoversættelse, hvorefter nogle prioriteringer og udfordringer i undervisning og forskning afslutter artiklen (afsnit 9).

2. Behov og muligheder

Sprogarbejde inden for en fagsproglig ramme kræver sproglig og faglig viden. Hvis den, der skal udføre arbejdet, ikke i tilstrækkelig grad besidder denne viden til at løse en stillet eller selvvalgt opgave, har hun behov for at indhente faglige og sproglige informationer, der skal transformeres til viden. Denne information kan med stor fordel søges i elektroniske kilder, ikke mindst online, hvor internet i dag er supermediet. I undervisningen gælder det både for underviser og studerende, at en lang række opgave-

typer kan løses helt eller delvist via få indtastninger på pc'en og klik med musen. Før semesterstart, her eksemplificeret ved den traditionelle face-to-face klasseundervisning, kan underviseren finde fag- og måske også genererelevante tekster (webdokumenter) på internet, som lovligt og nemt kan formidles via opgivelse af webadressen. Det er ikke kun tekster til løbende arbejde med det faglige stof, men også sprogligt orienterede opgaver og eksamenstekster, som underviseren med lidt finger-spitzengefühl (bl.a. kendskab til nøgleadresser på nettet inden for sit fagområde) samt evt. lidt efterfølgende dokumenttilpasning (typisk en reference til en given del af dokumentet) kan finde og referere til online. Den studerende kan via basalt kendskab til søgeteknikker og -værktøjer samt webadresser supplere de opgivne kilder med andre tekster. Både underviser og studerende kan udover nyttige fuldtekster i form af autentiske webdokumenter hente bearbejdet faglig og sproglig information, der af funktion minder om opslagsværker som encyklopædier og ordbøger. Alt det her beskrevne kan naturligvis også være relevant som bestanddel af andre undervisningsformer end de klassiske, ikke mindst virtuelle læringsstile.

3. Procestrin i Information Management

Informationsbehovene i forbindelse med sprogarbejdet tegner sig og tilfredsstillendes hensigtsmæssigt etapevis som funktion af de muligheder, internet tilbyder. Information Management (IM) udgør en ramme, der sekventielt beskriver processen efter opstået og formuleret behov, fra initial søgning til final præsentation af den løste opgave. De procestrin, det typisk kan dreje sig om, er alle centreret omkring information:

1. informationssøgning
2. informationsvurdering
3. informationsregistrering
4. informationspræsentation

Til hvert procestrin knyttes værktøjer og webressourcer, der introduceres i undervisningen, og som de studerende herefter arbejder systematisk med til løsning af opgivne eller selvvalgte opgaver af faglig og sproglig karakter. Det kan p.t. typisk dreje sig om:

- ad 1.: webbrowser (Microsoft Explorer) og søgemaskine(r) (Google og Yahoo! + metasøgemaskine: fx Copernic eller MetaCrawler)
- ad 2.: kvalitetsvurderingskriterier generelt
(fx Danmarks Biblioteksskole (se Skov 2002) eller program, der foretager konkret kvalitetsbedømmelse mod gebyr (fx Doctor HTML))
- ad 3.: databaseværktøj (fx Office-programmet Access fra Microsoft eller FileMaker Pro fra Lotus)
- ad 4.: præsentationsværktøj (fx Office-programmerne PowerPoint (grafisk præsentationsværktøj) og FrontPage (webeditor)).

Det kan være hensigtsmæssigt at udvide de fire processtrin med flere, men for afgrænsningens skyld behandles kun de fire i denne artikel, med vægt på informationssøgning. Fig. 1 viser en udbygget model taget fra et undervisningsmodul for sprogmedarbejdere under Åben Uddannelse på Handelshøjskolen i Århus.

4. Informationssøgning

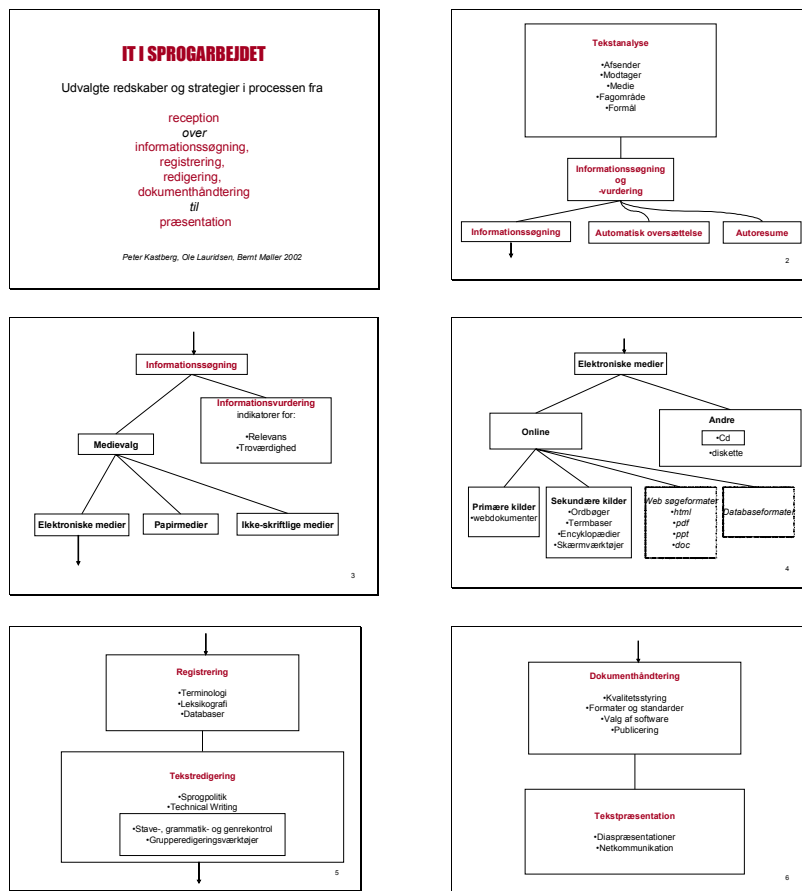
Når underviseren til baggrundslæsning skal finde fagtekster på webben til sine studerende eller disse selv skal finde tekster frem, startes med benævnelsen på det pågældende fagområde på de(t) relevante sprog. Hvis fagområdet ikke er specielt kulturspecifikt, kan søgninger på andre sprog også forsøges – typisk engelsk, der altid har oceaner af webmateriale. Metodisk set er der nu flere muligheder.

Hvis **webadressen (URL)** allerede er gemt som bookmark i **webbrowserens** mapper, skal den bare hentes frem fra *Foretrukne* i menuen i Microsoft Explorer (eller tilsvarende webbrowser).

Kendes webadressen allerede på en central kilde, indtastes denne i webbrowserens adressefelt, og der klikkes videre fra den pågældende (velkomst)side; en variant hertil er det hensigtsmæssige kvalificerede adressegæt. Vigtigheden af en kort og især gennemskuelig URL er gået op for snart alle aktører på nettet, og de fleste **internauter** ved, at en firmaadresse ender med enten *.dk* eller *.com* (evt. begge muligheder). Så skal man finde den danske vindmøllefabrikant Vestas på nettet, er den simpleste løsning at gætte på og taste *www.vestas.dk* eller *www.vestas.com* i browserens adressefelt. Vigtigt her er fuld opmærk-

somhed omkring ikke at lave tastefejl ved indtastning i URL-adressefeltet og samtidig at tænke på mulig alternativ stavning, typisk flertal vs. ental, forkortet form, engelsk stovevariant, ikke bruge æ, ø og å, unnlade blanktegn, mv. I modsat fald når man måske ikke frem. Nogle webmastere har dog sørget for alternative URL-varianter af samme årsag.

Fig. 1



Ønskes gode og indholdstunge, men ukendte adresser inden for fagområdet, er **søgemaskinevarianten katalog** som Yahoo!s faglige indeks sagen, da det viser en oversigt over udvalgte adresser, som man kan klikke sig videre fra. Her kan man også sammenligne de faglige hierarkiers benævnelser og adresseoversigt sprogene imellem (www.yahoo.dk vs. www.yahoo.fr).

Søgemaskinevarianten **robot**, som fx den p.t. altdominerende Google, bruges med stor fordel til at finde webdokumenter, der indeholder en specifik søgestreng, som indtastes. Til forskel fra kataloget søger robotten som default i hele teksten og ikke kun i teksternes overskrifter, men andre muligheder gives i *Avanceret søgning*, jf. afsnit 4.1.2. Når man, typisk efter at have læst om fagområdet i tekster fundet via en kendt URL, katalog eller robot, ønsker specifikke terminologiske og sproglige oplysninger i et bestemt dokument, hjælper Explorers søgefunktion *Søg (på denne side)* med at finde ønskede ord i form af tegnstreng, som tages i søgefeltet. Hvis man har søgt og fundet et dokument via Googles søgefelt kan man finde søgeordets placering i selve dokumentteksten ved i resultatoversigten at klikke på *cached* – søgeordet vil så fremstå omgivet af farve i teksten. Apropos farver: Google har en fremragende mulighed for at søge billeder, hvis tekst indeholder søgeordet. Denne facilitet er ikke bare stimulerende, men ofte en stor hjælp til at afkode indholdet af et ord, der lader sig afbilde.

Robotten giver en anden meget givtig mulighed for både underviser og studerende, nemlig attestation af sprogbrug via kvalificerede gæt. Det kan være eksistens, stavning, grammatik, konkordans og idiomatik, store eller små bogstaver – det meste kan tolkes ud fra antal returnerede resultater opgivet af søgemaskinen oven over listen med søgeresultater. Er der fx noget, der hedder en *lingvistik*? Hvordan staves *kom(m?)is(s?)ær?* Hedder det *en* eller *et parameter*? Lægges et dokument (*ud/op/–*) på (*internet/internettet/nettet*)?¹ Skrives *IT* med stort? Oplagt er også afgørelse af synonymers udbredelse opgjort som sammenlignede frekvenser for et givet søgesprog: er fx *lørner* mere eller mindre brugt på webben end *learner*? Det er en disciplin inden for søgemaskinebrug, som jeg og mine studerende har haft stor glæde af, selv om der naturligvis

¹ Dette afgøres måske bedst med robotten AltaVistas *proximity search* funktion NEAR.

ofte findes opslagsværker med lignende informationer. En ting er dog, hvad bøgerne normativt foreskriver, en anden hvad sprogbrugsrealiteterne er. Og det er ikke indlysende, at studerende i en demokratisk internetverden kun skal lære, hvad eksperterne har bestemt sig for er rigtigt. Det skal dog samtidigt understreges, at fejltolkninger inden for sprogtjek tit er overhængende realistiske – især er polysemi og homonymi farlige fætre, der kræver blik for alternative betydninger til ens søgeord.

Metasøgemaskinen udvider dækningsprocenten af webdokumenterne ret betragteligt uden dog at være sprogspecifik, hvilket reducerer nytteværdien noget. Metasøgemaskinen Copernic (gratis basisudgave) eller Metacrawler præsenteres i undervisningen til illustration af dækning af nettet via mulighed for valg af stort antal søgemaskiner, suppleret af konkret søgning, der viser hvilke søgemaskiner, der har fundet dokumenter med søgeordet.

To meget nyttige **search utilities** er Atomica og Alexa. Det er gratis et-kliks søgeværktøjer, der, efter download, med et enkelt højreklik fra skærmen finder en række informationer om søgeordet. Atomica viser ordbogs- og encyklopædisk stof, udtale (amerikansk), oversættelse (til hovedsprogene plus svensk, nederlandsk og græsk samt flere ikke europæiske sprog). Alexa giver adgang til kontaktfacts, børskurser, trafik til websteder med søgeordet, links hertil, webstedets historie og tidligere versioner, m.m. Autoriteten *Searchenginewatch*. *Search utilities* giver en desværre p.t. ikke helt funktionsduelig liste over 18 sådanne interessante værktøjer, men de små forklarende tekster siger en del om den fremtidige udvikling inden for søgeværktøjer.

4.1. Net, søgemaskine og tekst

Ved benyttelsen af ovennævnte muligheder har underviser og studerende brug for en bred og grundlæggende viden om nettet, webben, tekster og søgemaskiner med vægt på sproglige aspekter, så en hensigtsmæssig strategi kan lægges, og optimalt udbytte opnås. Nedenfor følger ret sum-

marisk en række punkter, som det vil være formålstjenligt at behandle i undervisningen.

4.1.1. Internetmediet

Begrebet **internet** bør i undervisningen præsenteres som benævnelsesmæssigt potentielt tvetydigt: internet ses ofte brugt snævert som synonymt med **webben**, men også, mere korrekt, som webben plus flere andre faciliteter, som fx e-mail og chat.

Afgrænsningen i forhold til det såkaldte dybe eller usynlige web (hvis sider typisk har endelsen *.asp*) er ligeledes vigtig, både fordi det dybe webs omfang er voldsomt meget større end webbens, og fordi dets indhold normalt er utilgængeligt for søgemaskiner.

En fornemmelse af antallet af tekster på webben og af hvor stor en del heraf, der kan nås via søgemaskiner, samt af forholdet mellem dokumenter på forskellige sprog hører også til i undervisningen.

Vigtigt er også, at de studerende gives en fornemmelse af genrer på nettet, et fænomen, der til dels afspejles i webdokumenternes format (mange virksomhedsgenrer findes ofte i pdf-format).

En generel diskussion af elektroniske kilder sammenholdt med papirmediet og traditionelle biblioteker er også på sin plads i undervisningen, såvel som en forklaring af karakteristikaene ved det multimediale (dvs. elektronisk kommunikation, der benytter sig af flere medieformer, som tekst, lyd, stillbillede, grafik, animation, ...), det interaktive (dvs. elektronisk kommunikationsform, hvor brugeren arbejder aktivt i dialog med programmet) og det dynamiske (typisk: løbende opdatering af webdokuments eller databases indhold).

Webbrowseren introduceres med specielt henblik på **URL-søgningen**, som alternativet til den søgemaskinebaserede søgning. Effektiv URL-søgning nødvendiggør endvidere en præsentation i undervisningen af URL'ens byggestene samt af top- og landomænenavne (*.com*, *.org*, ... *.dk*, *.fr*, ...) og af kvalificerede gættestrategier i den forbindelse.

4.1.2. Søgmaskinen

Alternativet til URL-søgning efter dokumenter fra specifikke adresser er som nævnt søgning vha. søgemaskine. Vigtigst p.t. er kendskab til Googles hjælpefunktioner *Avanceret søgning*, *Præferencer*, *Sprogværk-*

tøjer og *Tips til søgning* (se evt. Referencer, afsnit 10). I undervisningen gennemgås eller henvises detaljeret hertil, for her ligger alle de vigtige informationer, som det er et must for Googles brugere at kende. Fig. 2, Fig. 3 og Fig. 4 viser de tre hjælpesider, hvor de nyttige indstillinger foretages.

Fig. 2

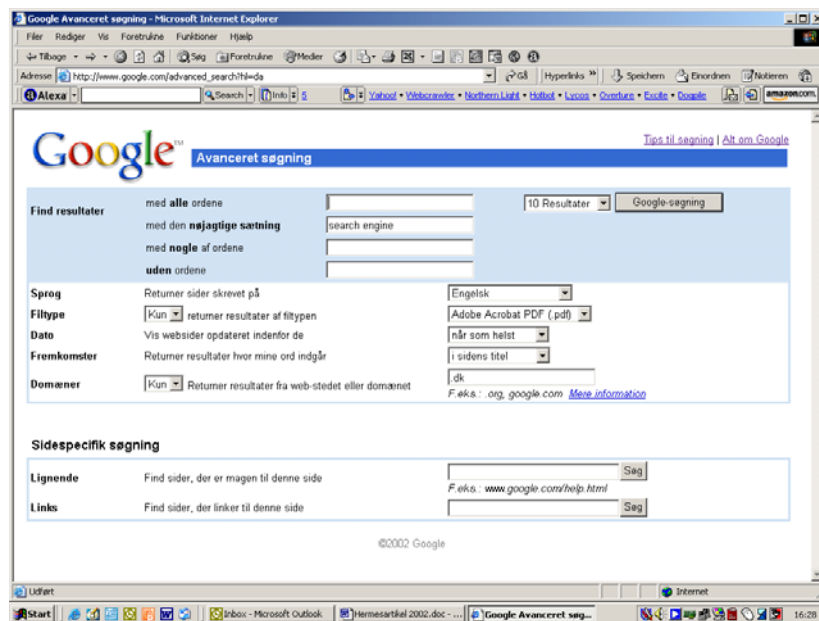


Fig. 3

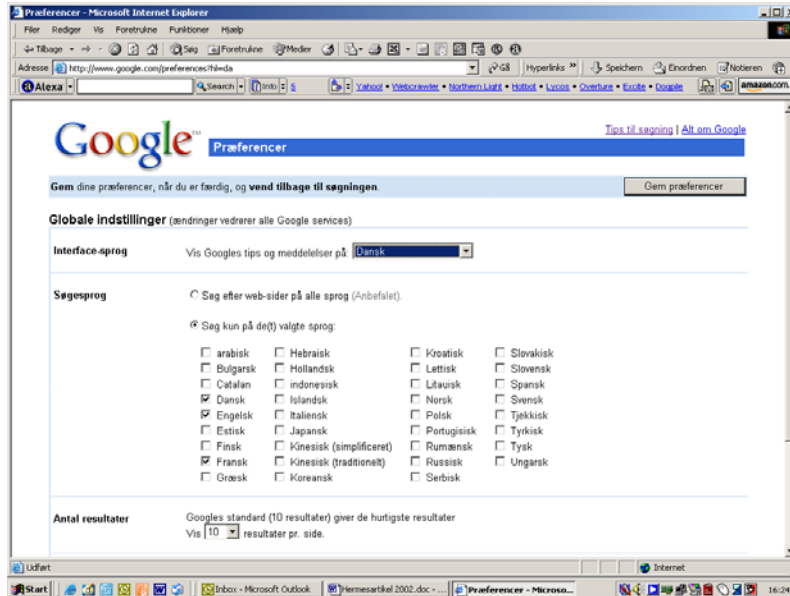


Fig. 4

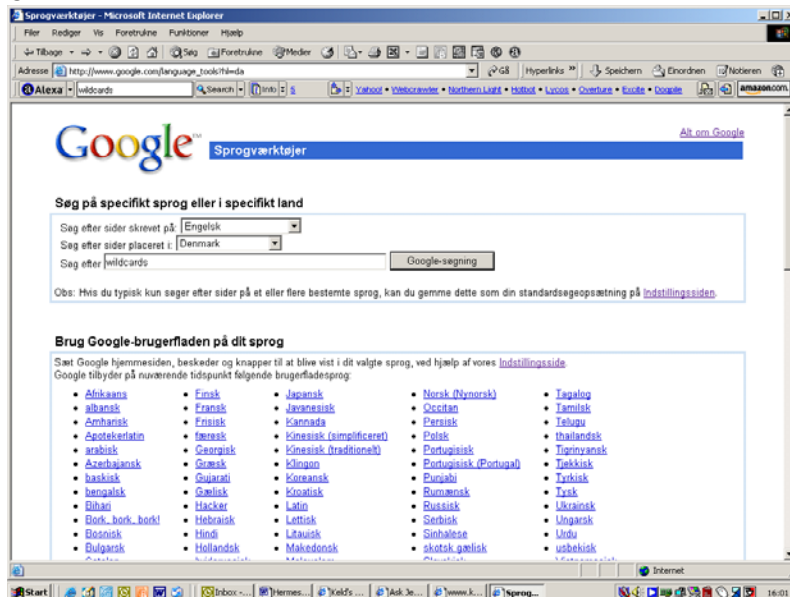


Fig. 5

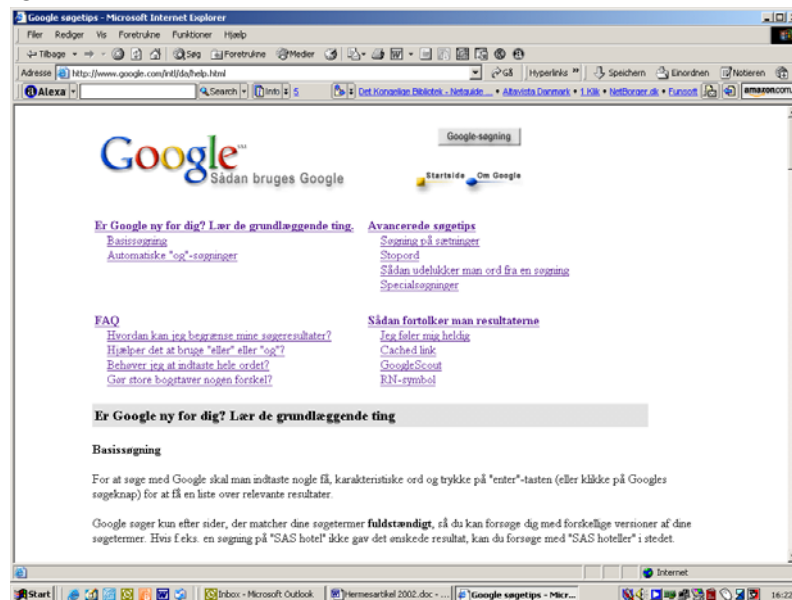


Fig. 2 (*Avanceret søgning*) tilbyder under *Find resultater* at vælge, om man for flere søgeord ønsker dokumenter 1) med alle søgeordene; 2) med den nøjagtige sætning; 3) med nogle af ordene; 4) uden ordene. Valg blandt disse gør en stor principiel forskel for de fundne dokumenter. For sammenlignende frekvensundersøgelser er valg af den nøjagtige sætning nødvendig, dvs. 2) oven for eller alternativt søgeordene omgivet af anførelstegn, ex ”*search engine*”. Avanceret søgning med Google giver mange flere muligheder, fx ønsket dokumentformat som fx word- eller pdf-dokumenter (*filtyper*), dokumentdel som fx titel, brødtekst eller URL (*forekomster*), domæner som fx *.dk* eller *.org.*, *lignende* (sider) og *links* (der henviser til den pågældende URL). Fig. 2 viser ordet *search engine* søgt i engelsksprogede pdf-dokumenters titel fra de dokumenter på danske webservere, der har domænet *.dk*. Et sådant dokument findes der p.t. et af, som Google kan indeksere.

Fig. 3 (*Præferencer*) viser de basale valgmuligheder af **interfacesprog** (som Google inkonsekvent også kalder brugerfladesprog) og af **søgesprog** – en skelnen, som man absolut bør bemærke sig. Figuren afbilder ønsket

om dokumenter på dansk, engelsk og fransk præsenteret med dansk-sproget interface.

Fig. 4 (*Sprogværktøjer*) giver muligheden for at søge i dokumenter på et valgt lands webservere. Eksemplet gengiver søgeordet *wildcards* søgt i engelsksprogede dokumenter placeret på danske webservere. Dem er der ca. 2.000 af – i øvrigt at sammenholde med ca. 750 i dansksprogede dokumenter (igen på danske webservere). Det er også under *Sprogværktøjer* man finder automatisk oversættelse af webdokumenter og webadresser på hovedsprogene, hvis man altså som interfacesprog har valgt engelsk, tysk, fransk, spansk, italiensk eller portugisisk. Som det fremgår, er sprog og internet ikke helt banale størrelser i dag – en fintunet søgning kan være hensigtsmæssig i flere tilfælde, hvor man ønsker overblik over nettets dokumenter og disses bestanddele.

Fig. 5 (*Sådan bruges Google*) fremkommer ved klik på *Tips* til søgning. Her kan man læse om grundlæggende og avanceret søgning i Googles setup samt alt om præsentation af søgeresultater og muligheder i den forbindelse. Nyttigt at vide om søgeprincipper er helt overordnet især, at Google søger præcis efter det, man skriver i søgefeltet og intet andet. Dette indebærer blandt andet, at man i modsætning til mange andre søgemaskiner og adgange til databaser ikke kan benytte sig af afkortede ordformer (**trunkering**) og ej heller symbol (som * og ?) for et eller flere vilkårlige tegn i søgeord, som man ellers kunne tænke sig for på en gang at søge fx efter entals- og flertalsformer (**wildcards**). Det indebærer også, at hvis man i søgefeltet skriver *prins felix* (med mellemrum) kommer ikke bare dokumenter om Prins Felix, men også dokumenter, hvor der et eller sted i teksten findes *prins* og måske et helt andet sted *felix*. Google viser dog øverst i resultatlisten de dokumenter, hvor afstanden mellem de to ord er mindst, typisk dem vedrørende Danmarks p.t. yngste prins.

I undervisningen kan også foretages sammenlignende søgninger mellem konkurrerende robotter (Google, Altavista, ...), så man kan se forskelle på fx søgeinterface og resultatliste – og selvfølgelig på antallet af hits for det aktuelle søgeord. Især er det instruktivt at sammenligne **resultatlistens hits**: om de først placerede hits er ca. de samme, om de omfatter et resume, om dette indeholder fremhævet søgeord, om de inkluderer URL, omfang, titel og dato for indeksering i søgemaskinens database – alt sammen brugervenlige indikatorer til afgørelse af værdien af det

pågældende hit, hvad enten man søger faglig information eller sproglig attestation.

Hvad sprogfolk derimod kunne savne, er redskaber til genresøgning, der kan indkredse typer af tekster af relevans for fx semiekspertes og lægmænd eller tekster som virksomhedsbrochuren, den populærvidenskabelige artikel eller manualen. I dag sker en sådan sortering oftest ad omveje eller vejindsnævninger som fx ved at begrænse sig til kendte virksomheder, til artikler fra bestemte tidsskrifter eller via søgning i bestemte filformater. Fremtiden ligger sandsynligvis i den såkaldte **intelligente agent**, der er internautens personlige digitale butler på nettet. Denne perfekte netbibliotekar varetager internautens interesser i forbindelse med fx informationsøgning og nyhedsopdatering samt fritidsarrangementer ud fra kendskab til faglige og personlige præferencer (Pedersen 1998). Et håb ligger også i **tekstmining**, en teknologi, der kan genkende grammatiske mønstre og finde sammenhænge mellem ordene i teksten. Matematiske modeller til tekstmining udvikles i disse år bl.a. til at klassificere og gruppere onlinetekster for eksempel i forbindelse med informationsøgning (Institute for Mathematics and its Applications). Man kunne forestille sig at disse avancerede teknologier også ville kunne bruges som et led i strategierne i den for oversættere vigtige jagt efter paralleltekster til supplement af hjælpen fra parallelt opbyggede emneportaler og flersprogede websteder og som alternativ til den meget chancebetonede udskiftning af nationalt topdomæne eller af anden sprogmærker i URL'en.

4.1.3. Webdokumentet

Gode søgemaskiner som Google dækker et meget stort antal websteder og er rimelig kvikke til at finde det ønskede. Især baseres udvælgelse og præsentationsrækkefølge (**ranking**) på linkstruktur og henvisninger, men også webdokumentets sproglige karakteristika i bred forstand indgår. Det drejer sig for det første om dokumentets basale karakteristika som skelnen mellem titel og brødtekst, hvor søgeord i titel som udgangspunkt kvalificerer til udtagelse eller en højere ranking blandt resultaterne end placering (kun) i brødteksten. Ligeledes er afstanden mellem flere søgeord i dokumentet (**proximity search**) et kriterium for rankingen, som beskrevet i afsnit 4.1.2. Interessant er det også, at dokumentets omfang i denne sammenhæng kan siges at være udvidet i forhold til en almindelig A-Z tekst, idet et søgeord, der indgår i URL'en eller i dokumentets links

også tæller i samme positive retning og som sådan er vigtigere end både et vilkårligt ord, men også et indholdsmæssigt vigtigt ord i teksten. Alt sammen ud fra hensynet til formodet relevans og tilgængelighed af de fundne dokumenter.

Et webdokuments sproglige opbygning er en disciplin i sig selv, som endnu kun er i sin vorden. Men enkle anbefalinger for en hjemmesides opbygning² og måske også for elektronisk kommunikation³ kan kort introduceres, idet de i første omgang kan bidrage til hurtigere afkodning af potentielt interessante webtekster. Den studerende vil derved bedre kunne forstå betydningen af links og utraditionelle læsemønstre⁴, af spillet mellem tekst, billede og lyd. Hun vil blive bevidstgjort om hvorfor mange videnskabelige og kommercielle webtekster er lidt anderledes tænkt og skruet sammen end tilsvarende papirbaserede og vil bemærke en afslappet, måske sjusket sprogbrug, der ofte befinder sig et sted mellem det mundtlige og det traditionelt skriftlige register. Hun vil også, uanset arbejdssprog, erfare de mange links til engelsksprogede dokumenter – et blandt flere vidnesbyrd om det engelske sprogs ret suveræne stilling i det globaliserede cyberspace.

5. Informationsvurdering

Efter informationssøgningen følger informationsvurderingen. Bruges URL-søgning, må det som regel forudsættes, at kvaliteten af kilden vurderes til at være i orden – man søger noget kendt. Eller gør man? Da URL'ene er blevet en konkurrenceparameter, et første middel til at fange internautens opmærksomhed ikke mindst i den meget store kommercielle del af webben, bruges en del energi på URL-design i virksomhedens eller institutionens branding. Her indgår givetvis genkendelighed og genkaldelighed som vigtige elementer. I det omfang denne strategi virker, vil informationssøgeren måske være fristet til at foretrække en godt markedsført URL i stedet for at ty til fx kataloget eller robotten. Ifølge Searchenginewatch.Search Utilities. viser undersøgelser, at kun 7– 8% af alle søgninger på nettet sker via søgemaskiner, 41% via **surfing**, dvs. videre-

² Om ikke lineære teksters disposition og byggekloster, se Jensen et al. 2001.

³ Om e-mail-sproget, se Caudery 2000.

⁴ **Hypertekst**, som er tekst, der indeholder (hyper)links til anden tekst inden for og/eller uden for det pågældende websted; hypertekst sigter på ikke lineær læsning.

klikninger fra andre webadresser, mens hovedparten i dag sker via **direkte navigation**, dvs. URL-søgninger (52% i 2002 mod 46% i 2001).

Sker informationssøgningen via et katalog som Yahoo!, er der på forhånd foretaget kvalitetsbedømmelse af Yahoo!'s stab. Tyr man derimod til robotten, er alt som udgangspunkt tilladt og muligt, og det er kun algoritmerne (de matematiske formler, der styrer indsamlings- og sorteringsprocesserne), der er på spil. Især i sidstnævnte tilfælde kan det være relevant at kigge nærmere på en række data i og omkring webdokumentet, specielt ved videnskabeligt og andet fagligt arbejde, hvor man bl.a. vil sikre relevans og troværdighed ved at undgå for store doser subjektivitet, kommercialisme og ikke ajourført materiale. Hvis man er usikker på dokumentets ophav, kunne et første kriterium være at tjekke URL'ens **domænenavn** efter en række kendetegn, fx for sandsynligt økonomisk sigte (.com) eller for enkeltmandskilde (tilstedeværelse af ~ i adressen). URL'en kan også afkortes bagfra for at se den underliggende adresse (normalt den samme som startside). Vigtig er kontaktmuligheden i form af e-mailadresse og angivelse af seneste ajourføringsdato. Og ikke mindst kildehenvisninger, herunder eksterne links, samt links fra andre kilder, der peger på webstedet. Dette sidste er jo netop hovedkriteriet, som Google anvender i sin ranking, hvilket giver høj sandsynlighed for relevans, men ikke nødvendigvis for kvalitet. Den sproglige kvalitet er et blandt flere elementer i det helhedsindtryk, man får ved bedømmelsen af kvalitet. Her kan meget andet også spille ind, fx reklamer og design og måske især funktionelle forhold som tilgængelighed og navigationsstrukturer (bl.a. site-map, frames, browser). Bibliotekerne informerer instruktivt om kvalitetskriterier, se fx Skov 2002 eller HHÅ's bibliotek.

Alt i alt er kvalitet på nettet en kompleks størrelse, der i mange sammenhænge måske ikke ofres den store opmærksomhed; man er tilfreds, hvis man har fundet noget emnemæssigt relevant på det pågældende sprog, eller hvis en mulig term eller ækvivalent kan attesteres. I mange sproglige arbejdssammenhænge er det dog min erfaring, at en kombination af kendskab til og beherskelse af (nogle af) ovennævnte redskaber og problemer, kombineret med sund fornuft, kritisk sans og lidt intuition, forhindrer fatale fejlskud.

6. Informationsregistrering

Når den ønskede information er fundet og kvalitetsbedømt, skal den registreres, sættes i system. Dette sker typisk via en simpel termbase (Microsoft Access eller den lidt mere brugervenlige, men mindre udbredte FileMaker Pro), enten i form af en færdigdesignet databaseapplikation, der udfyldes af de studerende, eller, mere instruktivt, men også krævende, i form af en tom databaseskal, der af den studerende designes til de typer informationer, hun begrundet ønsker. Da mine studerende skal applicere Terminologilærens principper for klassifikation og dokumentation af faglige relationer, ligger vægten på begrebsrelationer og ækvivalens. De obligatorisk udfyldte felter i databasen er *termækvivalent* og *definition*, de fakultativt udfyldte felter er fx *grammatik*, *eksempler*, *synonymer*, *kontekst*, *bemærkninger* (se eksempel Fig. 6). Indholdet i felterne skal støtte og fastholde det faglige aspekt.

Fig. 6

Et sekundært, men vigtigt delmål for termbasearbejdet, er at udnytte kompatibiliteten mellem i dette tilfælde databaseprogrammet og internet-

tets multimedialitet. Derfor opfordres de studerende til at søge sådanne relevante filer på nettet – fx lyd, grafik, billeder, animationer – og indsætte dem i deres nu multimediale termbase.

Et næste skridt kunne være at lægge **databasen** ud på internet. Andre muligheder kunne være at udvide konceptet til de lidt mere avancerede relationsdatabaser eller at give termbasen en mere leksikografisk vinkling. Registreringskonceptet kunne også skifte fokus fra termer og begreber til tekster og dokumenter, eller termdimensionen kunne suppleres hermed. Her vil internetbrowserens klassifikation af downloadede websteder i form af mapphierarkier kunne udbygges med grænsenedbrydende grafiske præsentationer af relaterede linksstrukturer, som fx programmet TheBrain tilbyder. TheBrain afbilder relationer mellem data – fx mapper, dokumenter og links – ”lige som hjernen gør det”, dvs. associativt eller ikke lineært hierarkisk ud fra et udgangspunkt, et centrum. Fig. 7 viser, hvorledes programmet ud fra sin webdatabase af 2,5 mio. URLs og 253.000 kategorier opfatter relationerne omkring kategorien *Translation*. Klikkes fx på *Tools* nederst til højre, flyttes denne kategori ind i centrum med nye relationer, der inkluderer *Translation* (Fig. 8), men omfatter nye kategorier, der ikke ses i Fig. 7.

Fig. 7

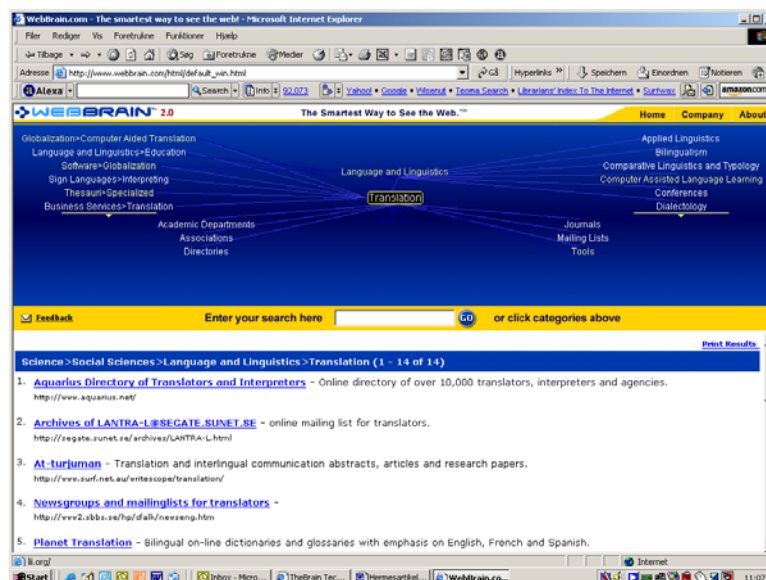
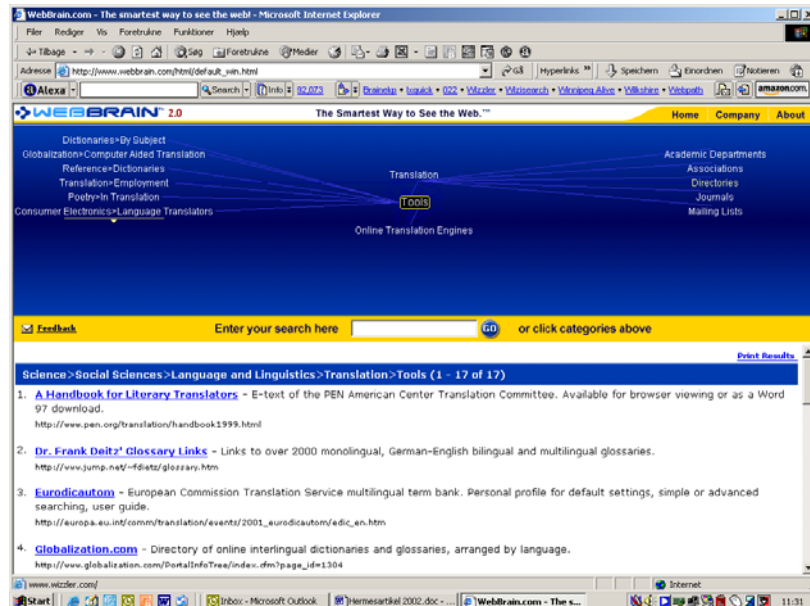


Fig. 8



Min erfaring fortæller, at termarbejde er en meget tidskrævende proces med opsøgning og vurdering af fagligt og sprogligt materiale på to sprog. Det bliver til gengæld en tilfredsstillende at se resultatet i elektronisk og systematiseret form med mulighed for multimedial integration og et tiltalende layout.

7. Informationspræsentation

Når informationen er systematiseret og registreret, kommer procestrinnet *informationspræsentation* på banen – medstuderende skal kunne få gavn af ens arbejde, også i elektronisk form. Den mundtlige fremlæggelse spiller imidlertid en rolle på hvert procestrin, også i skriftligt orienterede undervisningssammenhænge som fx oversættermoduler. Informations-søgningen og -vurderingen skal måske dokumenteres løbende på stor-skærm via internetopkobling, hvor de studerende klikker sig frem til gode valgte og til dårlige frasorterede websteder. Tilsvarende kan informationsregistreringen vises på basis af en databaseapplikation på computer eller netværk.

De to typer programmer, der oplagt kan refereres specifikt til under punktet informationspræsentation er præsentationsværktøj til computerstøttede **diaspræsentationer** og værktøj til redigering af et websted, **webeditoren**. I Microsofts kontorprogrammepakke Office hedder programmerne henholdsvis PowerPoint og FrontPage.

Mine studerende i teknologisk sprog anvender PowerPoint til faglige mundtlige præsentationer på klassen. Disse kan være rent factsorienterede, en redegørelse på fremmedsproget for dokumenteret centrale faglige og terminologiske sammenhænge, fx gensplejsning og økologi. Denne redegørelse baseres på computerdias kreeret ud fra en valgt skabelon eller fra grunden med eget design. Præsentationen kan også være en argumentativ fremstilling af et debateme, fx gensplejsede vs. økologiske madvarer. Begge former skal både integrere links til centrale websteder og til relevante sekundære onlinekilder som ordbøger og encyklopædier og i relevant omfang også til multimediefiler som lyd- og billedfiler.

Vores studerende på ÅU-modulet *IT i sprogarbejdet* (se Fig. 1) trænes bl.a. i Access, PowerPoint og FrontPage og skal i deres IT-eksamenprojekt med fokus på nogle af disse programmer demonstrere praktisk brugerkompetence og evne til planlægning, gennemførelse og præsentation af den selvvalgte opgave. Typisk laves et websted og/eller en database, der linkes til fra en PowerPoint applikation, det meste baseret på informations-søgnings- og -evalueringsstrategier behandlet i undervisningen.

Det forekommer mig indlysende, at studerende i almindelighed pga. disse præsentationsværktøjers utallige relevante muligheder og professionelle layout får boostet deres læring via en øget arbejdsglæde og derved motivation for at arbejde med det faglige/sproglige stof.

8. Ressourcer på nettet

Internet kan sammenlignes med et kæmpe bibliotek, hvor webstederne svarer til bøger og webdokumenterne tilnærmelsesvis til kapitler. Forskellen ligger ikke primært i størrelsen, men i graden af organisation af enhederne. Til forskel fra biblioteket er der ikke megen autoritet på nettet, enhver kan sætte sin egen bog ind på næsten selvvalgte hylder. Til at finde ”bogen” er der jo nettets bibliotekarer, som er webbrowsersen og søgemaskinen. Denne ”bog” er blot en af utallige ressourcer på nettet, der dækker i princippet alt, der kan hentes frem på skærmen via en web-

browser – rapporter, tidsskifter, lydfiler, bøger, databaser, programmer...you name IT. Inden for sproglig læring er disse ressourcer også relativt omfangsrige og på/om mange sprog. Den typologi, jeg vil anvende, er følgende:

1. Tekster
2. Portaler
3. Opslagsværker
4. Programmer

8.1. Tekster

Med tekster menes webdokumenter uden tematisk, sproglig eller filtypebegrænsning, og som ikke har portal-, opslags- eller programfunktion. De relaterer i denne artikel til de fleste afsnit og indgår i snæver forstand ikke i ressourcebegrebet.

8.2. Portaler

Portaler er oftest tematisk ordnede oversigter – typisk systematisk og/eller alfabetisk – over hovedsagligt eksterne links (dvs. links, der modsat interne links peger over i andre websteder). Sprogportaler kan dække delområder af det traditionelle almen- og fagsproglige univers og kan også indeholde links til websteder for biblioteker, arkiver, boghandlere osv., hvori man kan søge videre online. Det nok bedste eksempel på en allround **sprogportal** er Yourdictionary, der indeholder alle de online-ressourcer, man kan ønske sig, eller rettere, findes det, så findes det her: mængder af sprog og fagområder, grammatiker, morfologisk analyse, terminologi, automatisk oversættelse, søgemaskiner samt dagens ord, udrydningstruede sprog, diskussionsforum, andre portaler, ... og selvfølgelig alt i ordbøger og encyklopædier. Sprogportalene kan underinddeles efter indhold, formål og målgruppe.

Ordbogsportalen viser typisk en kortere eller længere linkliste over sproglige og evt. faglige opslagsværker som ordbøger og encyklopædier. Opslagsværkerne er oftest leksikalsk orienterede, men fx tekstkorpora kan også forekomme. Denne type sprogportaler har normalt til formål at give en samlet generel oversigt over ordbøger m.v. og sigter ikke på en specifik målgruppe. En ”god” ordbogsportal synes alt andet lige at have

et kvantitativt sigte – jo længere linkliste, desto større anseelse. Et fransk eksempel er Dictionnaires et lexiques.

Oversætterportalen er indholdsmæssigt bredere end ordbogsportalen. Den linker udover opslagsværker typisk til programmer til bl.a. online-oversættelse – gratis som kommercielle – til humanoversættere, tekstdatabaser, fagtidsskrifter, foreninger, jobs, ... og vægter nutidigt, praktisk og teknologisk indhold højt. Oversætterportalens formål er at assistere en bestemt profession – oversætterne – som er dens intenderede målgruppe. I modsætning til ordbogsportalen er indholdet ofte nøje udvalgt, så en rimelig kvalitet kan påregnes. Den mest autoritative oversætterportal er nok The Translator's Home Companion, der indeholder alt, hvad oversætteren måtte begære, eksempelvis formular til jobsøgning som oversætter.

Undervisningsportalen kan være en oversigt over en institutionsenheds undervisningsforløb, herunder e-learning, som fx HHÅ.Sprog, eller et websted med et konkret undervisningsmodul med bl.a. oversigter over sprogværktøjer som fx Gawinski 2002. Undervisningsportalens funktion er at give målgruppen en samlet adgang til alle de informationer, der kan formidles elektronisk. I endnu højere grad end oversætterportalen er indholdet meget specifikt rettet mod en veldefineret målgruppe – grupper af sprogstuderende – og derfor snævrere. En god undervisningsportal er kendetegnet ved en meget høj grad af bevidst udvælgelse af interne og eksterne links og andre oftest praktiske eller direkte informationer, som opgaveformuleringer og lektionsplaner.

8.3. Opslagsværker

Opslagsværker dækker over websteder, der er konciperet hovedsagligt til sproglige opslag via søgefelt i egen database og/eller surfing (klikning) i alfabetisk/systematisk oversigt, der modsat portalerne overvejende består af interne links. De kan omfatte, hvad der benævnes eller svarer til **ordbøger**, ordlister, glossarer, leksika, terminologiske lister mv., der overordnet set tilnærmelsesvist er synonyme. Et eksempel på en gratis online flersprogsordbog (ty–en–fr–nl) er Wörterbuch Wortschatz. Den udmærker sig ved sine søgbare og meget omfangsrige tekstkorpora på engelsk, tysk, fransk og nederlandsk, som giver informationer om bl.a. kontekst og

kollokationer. Her kan også nævnes det danske Korpus 2000, der udover kontekst også viser diakrone data (1990 → 2000).

Gruppen omfatter også **encyklopædier** og andre mere fagligt orienterede opslagsværker samt termbaser/termbanker. Et eksempel på en encyklopædi er Microsofts Encarta, der omfatter en gratis ordbog og en delvis gratis encyklopædi med multimediereferencer til ”alt” fra et leksikon-univers.

Metaordbøger minder om portaler; de søger simultant i flere ordbøger og evt. andre opslagsværker. Det klassiske eksempel er den flersprogede OneLook Dictionary, der omfatter over 800 ordbøger med 5 mio. opslagsord (ca. 700 med engelske, 90 med franske, 60 med tyske, 50 med spanske, 30 med italienske og 13 med kinesiske opslagsord), hvor de respektive sprog indgår også inden for fagområder som teknologi, naturvidenskab, medicin, økonomi og IT. Søgeresultaterne vises som linklister med præsentation af de ordbøger, søgeordet findes i. Der er mulighed for at søge sprog- og fagområde for sig samt specifikt definitioner og oversættelse. Hertil kommer den nyttige facilitet, der vises for hvert opslag: *Phrases that include the word X*.

Termbaser/termbanker er ofte utilgængelige for uindviede (passwordbeskyttede). Kendte gratis internationale termbaser omfatter EU’s Eurodicautom, der p.t. synes at afvente en europæisk superdatabase (IATE). Eurodicautoms værdi i oversættelsesarbejdet er betragtelig, idet termbasen omfatter EU-kommissionens meget omfattende terminologisamlinger, der inkluderer dansk i væsentligt omfang. Le Grand Dictionnaire Terminologique er en enestående, gigantisk og gratis fransk–engelsk termbase udarbejdet af sprogmyndighederne i Québec. Ud fra et meget simpelt søgeinterface vises tematiske indeks, hvorfra mere præcise termer kan vælges. Kvalitet og brugervenlighed af resultatpræsentationen er i absolut topklasse med angivelse af ækvivalent, synonymer, kvasisynonymer på begge sprog, definition, detaljerede kommentarer om bl.a. termdannelse og (canadiske) anbefalinger.

8.4. Programmer

De ressourcer på nettet, som er tiltænkt en sproglig systematisering, udvælgelse eller transformation af tekster eller tekstdele på nettet falder ind under kategorien *programmer*. Det kan fx være programmer til onlineoversættelse eller konkordansprogrammer.

Programmer til onlineoversættelse af webdokumenter og indtastet tekst er stadig på et ret rudimentært niveau, selv om forbedringer kan konstateres løbende. Dansk af brugbar kvalitet findes ikke, men mellem hovedsprogene er der flere oversættelsesprogrammer, bl.a. Reverso og Systran. Begge programmer tilbyder gratis onlineoversættelse af indtastet tekst og af hele webdokumenter mellem engelsk, tysk, fransk og spansk. Systran oversætter endvidere fra/til italiensk, portugisisk, russisk, koreansk, japansk og kinesisk. Disse to programmer har vi i oversættelsestimerne kåret til de bedste/mindst ringe på basis af sammenlignede oversættelser af dele af webdokumenter omhandlende IT fra/til fransk og engelsk, herunder tilbageoversættelser. Denne øvelse krydres med oversættelse af anekdotiske engelske sætninger som *My name is Bond, James Bond*, hvor *James* i et (dårligere) program for 1–2 år siden blev oversat til norsk som *syltetøy*, mens *Bond* stadig i flere programmer oversættes, som om det var en velkendt type værdipapir.

Konkordansprogrammer kan downloades fra nettet, men de fleste synes primært beregnet til offline tekstkorpora og ikke til online webdokumenter. Ønsker man at bruge webbens enorme tekstmateriale som korpus, kan WebCorp dog p.t. anbefales. Dette korpuslingvistiske program finder og gengiver brugervenligt de kontekster, som søgeordet/søgeordene er omgivet af. Programmet tilbyder også en række værktøjer, vi kender fra søgemaskinerne, såsom domænespecifik søgning, wildcards og ordfilter, men man kan ikke direkte vælge søgesprog. Det majsommelige alternativ til konkordansprogrammer er at downloade et udvalg af tekster (se herom Austermühl 2001, kap. 8).

9. Perspektiver

Internet som tekstsamling er en potentiel guldmine til sproglig læring for både undervisere og studerende. Der kommer stadig flere relevante dokumenter til på langt de fleste sprog, dokumenter som alt i alt bliver stadig lettere at finde, både hvis man søger på et emne, og hvis man ønsker et bestemt dokument. Det forekommer mig, at hurdlen for internauten knap så meget ligger i kvaliteten af det fundne – især takket være søgemaskinens ranking – som i de endnu begrænsede muligheder for at præsentere sin egen søgeprofil og søge på genrer, på dokumenters tekstuelle karakteristika. Men intelligente agenter og tekstmining bliver banale værktøjer en dag, ligesom informationssøgning, -vurdering og -præ-

sentation i de allerseneste år tyder på at være ved at blive det i læringen. Imidlertid er tekster ikke kun hele tekster, der handler om noget og er præsenteret på en bestemt måde. Tekster indeholder tegnstrenger, bogstaver og tal især, der er en markant dokumentation af usus inden for kommunikation beregnet til skærmlæsning eller til udprintning fra et elektronisk medie. I et større perspektiv er disse tekster en stadig vigtigere manifestation af skriftsproget i iøjnefaldende udvikling. Denne deskriptive skattekasse af ord og sætninger burde udnyttes meget mere i sprogarbejde af alle typer, hvilket søgemaskinerne, som beskrevet, giver visse muligheder for, men faciliteter som fx proximity search og konkordans er endnu ikke legio. Sprogforskningen har en del at byde ind på, blandt meget andet en teknologisk betinget ajourføring af tekstbegrebet samt en beskrivelse af hypertexters og webgenrers strukturer og virkemidler samt disses betydning for tekstredigering og for læsemønstre, videntilegnelse og læringsprocesser.

Internet som ressourcebank er tilsvarende meget værdifuld. Nogle ressourcer har metakarakter i forhold til webdokumenterne, fx portaler og programmer til onlineoversættelse samt visse opslagsværker, der baserer sig eller skal anvendes på informationer hentet fra webdokumenter. Andre ressourcer har metakarakter i forhold til informationer, der går ud over tekster på nettet, som det er tilfældet for mange onlineordbøger og encyklopædier. Uanset metastatus venter der endnu en forskningsindsats for alle kategorier af ressourcer. Ligesom undervisningsportalen må formodes kun at indeholde links til anbefalelsesværdigt materiale, burde ikke mindst ordbogsportaler ikke lade den besøgende i tvivl om indhold og status af et givet link. Her kan en lille metatekst, fx i form af en såkaldt **mouse-over** label, hjælpe på vej, helst også tilknyttet en kvalitetsvurdering. Forskning i onlineordbøger (og andre sproglige/faglige opslagsværker) er endnu kun i sin vorden og må nødvendigvis tage udgangspunkt ikke bare i metaleksikografien, men også i discipliner, der har fokus på digitalisering, multimedialitet og ikke mindst interaktivitet og hypertext.

Min erfaring tilsiger mig, at en prioritering blandt IM-etaper har det med at blive påtrængende nødvendig. Man bør dog erindre, at dette til dels opvejes af den stigende grad af interface-brugervenlighed og, som det hedder, de studerendes **computer literacy** og **hands-on competence**. Vigtigt er det, at man som underviser er opmærksom på, i hvilket omfang de studerende kan interagere målrettet og udbytterigt med interfacerne.

Men nok så væsentligt er det, at man sikrer sig, at de har et vist overblik over internet i teknologisk og især sproglig kontekst samt over proportioner og basal logik. Uden disse kompetencer risikerer informationsøgningen, resultatudvælgelsen og tolkningen af det fundne at blive for vilkårlig og kvalitativt tynd. I informationsøgningen bør man også prioritere oplæring i at foretage kvalificerede valg og gæt. Dette gælder både med hensyn til kodning og afkodning af webadresser, mht. til valg blandt alternativer i søgemaskinernes resultatlister og ikke mindst mht. til valg af formulering og sproglig–grammatisk form til indtastning i søgefeltene, især selvfølgelig hvis formålet er sproglig attestation. I informationsregistreringen og -præsentationen bør samspillet mellem software og internet indgå som en rettesnor for IM, så maksimal integration bliver et delmål i opgaveløsningen.

Webdokumenters adresser kan enten opgives af underviser til studerende eller omvendt. Selv foretrækker jeg som regel i mine timer i teknologisk oversættelse en kombination: jeg udleverer fx de franske adresser og mine studerende skal så på nettet finde, kvalitets- og relevansvurdere danske paralleltekster på nettet, der matcher de franske, primært mht. emne og sværhedsgrad. Hvad angår de mere værktøjsprægede ressourcer, der kan komme til at spille en nøglerolle i opgaveløsningen – portaler, ordbøger, encyklopædier og databaser – videreformidler jeg dem fra starten af sprogundervisningen, der naturligvis bør foregå i et lokale med pc, internet, projektor og storskærm, så kombinationen af *IT og sprog* er til at få øje på – også helt bogstaveligt.

Referencer

- Alexa. <<http://www.alex.com>> (26.08.2002).
- Arleth, Nina 2001: IT og sprog – få øje på kombinationen. In *Computerworld* 11.12.2001.
- Atomica. <<http://www.atomica.com>> (26.08.2002).
- Austermühl, Frank 2001: *Electronic Tools for Translators*. St. Jerome Publishing.
- Caudery, Tim 2000: *E-mail og e-mailsprog*. <<http://www.dpb.dpu.dk/infodok/sprogforum/spr16/caudery.html>> (26.08.2002).
- Copernic. <<http://www.copernic.com>> (23.08.2002).
- Dictionnaires et lexiques. <<http://www.swarthmore.edu/Humanities/clicnet/dictionnaires.html>> (23.08.2002).
- Doctor HTML. <<http://www2.imagiware.com/RxHTML>> (26.08.2002).
- Encarta. <<http://encarta.msn.com/>> (23.08.2002).

- EURODICAUTOM. <<http://europa.eu.int/eurodicautom/login.jsp>> (23.08.2002).
- Gawinski, Birthe. <<http://esb.sdu.dk/birthe/ord.html>> (23.08.2002).
- Google. <http://www.google.com/advanced_search> ; <<http://www.google.com/preferences>> ; <http://www.google.com/language_tools> ; <<http://www.google.com/intl/da/help.html>> (23.08.2002).
- HHÅ's bibliotek. <<http://www.hba.dk/metro>> (30.08.2002).
- HHÅ. Sprog. <<http://www.sprog.asb.dk/undervisning.htm>> (29.10.2002).
- Institute for Mathematics and its Applications. <<http://www.ima.umn.edu>> (29.10.2002).
- Jensen, Ulf Joel et al. 2001: *skrivgodt.dk*. Frydenlund/Grandjean kommunikation.
- Korpus 2000. <<http://korpus.dsl.dk/korpus2000>> (29.10.2002).
- Le Grand Dictionnaire Terminologique. <<http://www.granddictionnaire.com>> (23.08.2002).
- Marriam—Webster Collegiate Dictionary. <<http://m—w.com/>> (23.08.2002).
- Metacrawler. <<http://www.metacrawler.com>> (23.08.2002).
- OneLook Dictionary. <<http://www.onelook.com>> (23.08.2002).
- Pedersen, Karim 1998: Intelligente agenter og Internet. <http://ix.db.dk/inettema/rapport98/06_kp.htm> (29.10.2002).
- Reverso. <<http://www.reverso.net>> (23.08.2002).
- Searchenginewatch. Direct navigation. <<http://www.searchenginewatch.com/sereport/02/02—nav.html>> (26.08.2002).
- Searchenginewatch. Search utilities. <http://searchenginewatch.com/links/Search_Uilities/Other_Search_Uilities/index.html> (23.08.2002).
- Skov, Annette 2002: *Kvalitetsvurdering af websider*. <<http://www.db.dk/bib/tutorials/kvalitet/default.htm>> (26.08.2002).
- Systran. <<http://babelfish.altavista.com>> ; alternativt <<http://translate.google.com/translate>> (23.08.2002).
- The Translator's Home Companion. <<http://www.lai.com/lai/companion.html>> (23.08.2002).
- TheBrain. <<http://www.thebrain.com>> (23.08.2002).
- WebCorp. <<http://www.webcorp.org.uk>> (30.01.2003).
- Wörterbuch Wortschatz. <http://wortschatz.uni—leipzig.de/index_js.html> (23.08.2002).
- Yahoo!. <<http://www.yahoo.com>> ; <<http://www.yahoo.fr>> ; <<http://www.yahoo.dk>> (26.08.2002).
- Yourdictionary. <<http://www.yourdictionary.com>> (23.08.2002).