

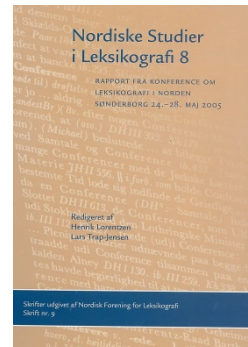
# NORDISKE STUDIER I LEKSIKOGRAFI

Titel: Nyt redigeringsystem – overvejelser og valg

Forfatter: Marie Bilde Rasmussen

Kilde: Nordiska Studier i Leksikografi 8, 2006, s. 347-358  
Rapport från Konferens om lexikografi i Norden, Göteborg 27.-29. maj 1999

URL: <http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive>



© Nordisk forening for leksikografi

## Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

## Søgbarhed

Artiklerne i de ældre Nordiske studier i leksikografi (1-5) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

## Nyt redigeringsystem – overvejelser og valg

**How to select a new dictionary writing system.** Before introducing a new dictionary writing system it is important to describe the functionality of the writing system and what it requires and needs. These requirements and needs are normally described in a so-called list of requirements. The list of requirements is used to test the dictionary writing systems that are under consideration. In the Dictionary Department at Gyldendal Publishers we have prepared an extensive list of requirements listing more than 450 items. In this article I would like to pass on the experience from our tests. Firstly, I offer a discussion of the concept of a dictionary writing system. Then, I examine various dilemmas and conditions and in each case describe the prioritisation they may cause.

I de seneste år har adskillige ordbogsproducerende institutioner i de nordiske lande diskuteret ordbogsredigering. I den forbindelse er markedet for digitale værktøjer til netop redigering og produktion af ordbøger blevet undersøgt. I en årrække har der slet ikke været nogen til salg, men nu er der igen grøde: I dag kan man vælge mellem flere systemer,<sup>1</sup> hvis man da ikke vælger at udvikle sin egen løsning.<sup>2</sup>

I forlaget Gyldendals ordbogsafdeling har vi længe været på udkig efter et nyt system til at afløse vores to gamle systemer, GestorLEX<sup>3</sup> og Compulexis.<sup>4</sup> I den forbindelse har vi udarbejdet en omfattende kravspecifikation med flere end 450 punkter. I kravspecifikationen har vi prioriteret de krav til funktionalitet og systemarkitektur mv., som vi har fundet passende at stille til et system, der skal opfylde lige netop Gyldendals behov nu – og forhåbentlig også i fremtiden. Det er med kravspecifikationen i hånden, at vi har kunnet afprøve forskellige systemer på en så grundig og systematisk måde, at vi har følt os overbeviste om at være kommet “langt ud i krogene” og “hele vejen rundt”, før vi har sagt ja eller nej til det enkelte redigeringsystem.

Det er erfaringerne fra dette arbejde, jeg gerne vil bringe videre i denne artikel. Det er derfor ikke en videnskabelig artikel – det er snarere en opsummering af, hvad vi lærte af det konkrete projekt. Alle ordbogsprojekter er forskellige, og det er institutioner og projekters organisation også. Derfor har jeg prøvet at generalisere mine overvejelser så meget, at de fremstår så relevante som muligt for andre, der står i samme situation, som vi gjorde.

---

<sup>1</sup> Jeg har i skrivende stund kendskab til *Dictionary Publishing System (DPS)* fra franske IDM, *Dictionary Workbench* fra tyske Acolada, *iLEX* fra danske EMP og *TshwaneLex* fra det belgisk/sydafrikanske TshwaneDJe.

<sup>2</sup> Som man har valgt at gøre det hos bl.a. Norsk Ordbok.

<sup>3</sup> Udviklet af danske TEXTware A/S.

<sup>4</sup> Udviklet af engelske Compulexis Ltd.

## 1. Hvorfor indfører man et nyt redigeringsystem?

Der kan være mange grunde til, at man ønsker at indføre et nyt redigeringsystem, men det er mit indtryk, at de fleste af disse kan sammenfattes i følgende:

1. For at opnå økonomiske fordele, fx fordi redigering og produktion kan effektiviseres
2. For at kunne genbruge og udveksle ordbogsdata i forbindelse med nye ordbøger
3. Fordi systemet giver mulighed for bedre redaktionel kontrol med ordbogens indhold
4. Fordi systemet kan yde redaktionel støtte til ordbogsforfattere og -redaktører
5. For at kunne søge på kryds og tværs i alle ordbogsdata

## 2. Hvad er et ordbogsredigeringsystem?

Et ordbogsredigeringsystem er et stykke software, der er udviklet med henblik på at redigere og producere ordbøger. I programmet til den 3. internationale workshop om ordbogsredigeringsystemer, der afholdtes på Mazaryk Universitet i Brno i september 2004, defineres termen *Dictionary Writing System* (kaldet *DWS*) således:

A dictionary writing system is a piece of software for writing and producing a dictionary. It might include an editor, a database, a web interface and various management tools (for allocating work etc.). It operates with a dictionary grammar, which specifies the structure of the dictionary. (DWS 2004)

Denne definition giver et udmærket indtryk af, hvilke komponenter der kan være integrerede i et ordbogsredigeringsystem. Men definitionen giver ingen beskrivelse af systemets funktionalitet bortset fra kravet om, at der skal indgå en ordbogsgrammatik, der fastlægger reglerne for ordbogens struktur.

Bo Svensén går anderledes konkret til værks i 2. udgave af *Handbok i lexikografi* (Svensén 2004:498). Her opremser han en række konkrete krav til den funktionalitet, et redigeringsystem skal have, bl.a. alfabetisk sortering, håndtering af homografi, søge- og statistikfunktion og kontrol af krydshenvisninger.

## 3. Hvad er et *generelt* ordbogsredigeringsystem?

Oftentimes hører man ønsket om et 'generelt system' fremsat. Det er dog ikke altid ganske klart, hvad der menes med det. Man kan fortolke *generelt* på mindst tre måder, når det gælder redigeringsystemer:

1. Systemet kan være bygget op af standardiseret software. Og data kan produceres eller (som minimum) eksporteres i et internationalt standardiseret

dataformat. Det generelle består da i, at både redigeringsystem og data kan tilgås i andre omgivelser end de specifikke.

2. Systemet kan indeholde netop én ‘generel’ ordbogsgrammatik. Det underforstås da, at alle ordbøger kan beskrives i én og samme grammatik. Her er det selve ordbogsbegrebet, der opfattes generelt.
3. Systemet kan være generelt i den forstand, at det tillader brugeren selv at definere ordbogsgrammatikker, præsentations- og produktionsrutiner til de enkelte ordbøger, der således ikke postuleres at være ens. Her består generaliteten i at kunne håndtere forskellige ordbøger.

### ***3.1. Systemet er bygget op af standardsoftware, og data lagres i standardformat***

Disse meget tekniske krav tilgodeser de ydre processer ved ordbogsarbejdet. Drift og vedligeholdelse gøres sandsynligvis både lettere og billigere ved brug af standardsoftwarekomponenter, og man er mindre afhængig af en bestemt leverandør. I nogle institutioner vil det være et krav, at alt bygges op af open source-software, og at systemet kan køre på uafhængige (og måske gratis) platforme som fx Linux.

At ordbogsdata lagres (eller i det mindste kan eksporteres) i et standardiseret, platformsuafhængigt dataformat, er et vigtigt krav, hvis man ønsker at kunne producere sine ordbøger billigere og mere effektivt. Også her bliver man mindre afhængig af bestemte leverandører. På samme måde bliver udveksling af data med andre ordbogsproducerende institutioner naturligvis lettere ved brug af et standardformat.

Disse krav er i sig selv relevante og efter min mening vigtige. Men de er samtidig utilstrækkelige, idet man slet ikke fokuserer på selve redigeringsprocessen eller på håndteringen af den leksikografiske struktur i en ordbog. I et ordbogsprojekt fylder både formulering af de redaktionelle principper og især selve redigeringsfasen naturligt nok langt det meste af forløbet.

### ***3.2. Systemet kan håndtere netop én ordbogsgrammatik***

Hvis redigeringsystemet kun opererer med én ordbogsgrammatik, som alle ordbøger i systemet skal følge, kan man effektivisere redigeringsproces og produktion ved at bruge prædefinerede rutiner til præsentation af data på skærmen (både til indtastning og preview) og ikke mindst ved at have faste rutiner til satsgenerering og anden publicering (fx transformation af data til html-format). En anden fordel er, at alle brugere altid skal lære nøjagtig denne ene ordbogsgrammatik at kende.

At udforme en sådan fælles grammatik kan ikke være nemt, hvis den skal gælde både mono- og bilingvale ordbøger (med forskellige målgrupper og ordbogsfunktioner), historiske ordbøger, begrebsordbøger m.fl. – måske også mere encyklopædiske opslagsværker.

En fælles ordbogsgrammatik kan være ‘maksimal’ eller ‘minimal’. En maksimal

fællesgrammatik vil indeholde så mange forskellige informationstyper som muligt i alle tænkelige strukturelle kontekster. Grunden til at vælge denne løsning vil være, at grammatikken rummer alle de informationstyper og kombinationer af disse, som man kan tænkes at få brug for i bare én af de ordbøger, der redigeres i systemet. Man vil helst ikke risikere at hæmme forfatterens udtryksmuligheder på nogen måde. Brug af en maksimal grammatik vil typisk bevirke, at systemet i en given strukturel kontekst tilbyder brugeren alt for mange (og for den enkelte ordbogstype ofte irrelevante) informationstyper. Der er kort sagt ikke megen redaktionel kontrol eller støtte at hente i denne model, og brugeren risikerer at drukne i de mange muligheder.

Arbejder man derimod med en minimal fællesgrammatik, går øvelsen ud på at beskrive den mindst mulige fællesmængde. Minimalgrammatikken vil ofte opleves som for restriktiv i den enkelte ordbog. Brugeren har jo behov for redaktionel støtte i form af tilbud om netop de relevante informationstyper i en given kontekst. Der er stor risiko for, at man overser nogle af de nødvendige strukturelle særtilfælde, der findes i enhver ordbog. En mulig konsekvens er, at den samme informationstype bruges på forskellig vis i to forskellige ordbøger. Det går ud over den redaktionelle kontrol og den nok så vigtige fælles forståelse. Desuden kan det give problemer i forbindelse med præsentationen af data.

Særlig grelt vil det være, hvis den fælles 'generelle' ordbogsgrammatik er givet på forhånd af redigeringsystemet. Så har man ikke mulighed for selv at fastlægge den leksikografiske struktur. Ordbogsgrammatikken skal nemlig være fælles for mange forskellige ordbøger.

### *3.3. Systemet kan operere med forskellige ordbogsgrammatikker til forskellige ordbøger*

Hvis redigeringsystemet er generelt i den forstand, at det tillader brugeren selv at formulere ordbogsgrammatikken og andre definitioner til hvert enkelt projekt, kan man opnå en meget høj grad af redaktionel kontrol og støtte. Man kan definere det enkelte projekt uden at skulle tage hensyn til andre værkers opbygning. Man kan koncentrere sig om at tilpasse den enkelte ordbogsgrammatik til den ønskede mikrostruktur og til redigeringsprocessen for det enkelte ordbogsprojekt. Det er en fordel, hvis man arbejder med flere ordbogsprojekter i huset. Det er også en god forudsætning for, at systemet kan få en vis udbredelse, dvs. en større kundekreds. Det sidste kan vise sig at være meget vigtigt også for den enkelte kunde, middleleveltiden for ordbogsredigeringsystemer taget i betragtning.

Med denne type system skal man bruge flere ressourcer på oplæring. Ud over at lære selve systemet at kende skal brugerne ikke længere kun lære at bruge én enkelt fællesstruktur og ét fællesinventar af informationstyper. For hver ordbog, de skal arbejde med, skal de sætte sig ind i et specifikt inventar og en specifik struktur. Hvis

ordbogsgrammatikken er tilrettelagt rigtigt, får brugeren til gengæld hjælp i form af støtte og stimulation til at følge de redaktionelle retningslinjer. Desuden vil det være nyttigt, hvis man kan tilpasse skærmpresentationen til indtastning til det enkelte projekt – måske endda til den enkelte bruger.

Det kræver ressourcer, når man selv skal udarbejde ordbogsgrammatik og præsentationsdefinitioner (til skærm og egentlig produktion). Men hvis man dermed både kan spare tid og øge datakvaliteten, er det værd at overveje.

Stordriftsfordele i form af genbrug af fx præsentations- og produktionsrutiner (hvor det giver mening) bliver ved valg af denne type system et redaktionelt ansvar i stedet for en indbygget egenskab ved systemet. Gevinsten er, at man får mulighed for selv at træffe leksikografiske, redaktionelle, pædagogiske og æstetisk-grafiske valg.

### 3.4. *Alternativ: at udvikle sit eget system*

Daniel Ridings fra enheden for digital dokumentation ved Oslo Universitet taler i Ridings 2003 for den stik modsatte løsning, nemlig at man i den enkelte institution udvikler et specifikt redigeringsystem pr. ordbogsprojekt. Han argumenterer for, at ethvert ordbogsprojekt bør have en softwareudvikler tilknyttet:

It is extremely unlikely that a project will find a commercial product that will meet its needs off the shelf. For software to be beneficial to lexicographers, it must be linked to the style manual of the relevant project. This makes the software user-friendly, but not commercially viable since it is specific, not generic. It is difficult, if not impossible, to find two dictionaries from two independent groups that have the same format. The software developer is a crucial member of the team of a dictionary project. (Ridings 2003:214)

Termerne *style manual* og *format* synes begge at dække over både ordbogsgrammatik og regler for ordbogens præsentation. Desuden antages det, at alle kommercielle systemer er af den type, der opererer med én fælles ordbogsgrammatik for alle værker. Endelig skelnes der ikke klart mellem programmering af selve softwaren (som typisk skal udføres af en programmør) og konfigurering af et givent ordbogsprojekt i form af udarbejdelse af ordbogsgrammatik, præsentationsdefinition mv.

Som det siges i artiklen, er specialudviklede systemer som regel kun meget lidt generelle. Det er selvfølgelig en fordel i den forstand, at man kan tilpasse systemet nøjagtigt til det enkelte ordbogsprojekt. Men man skal også kunne ændre systemet, hvis man midt i et projektføreløb ønsker at lave om i strukturen. Hvis man har flere ordbogsprojekter, kan det være svært at genbruge systemet projekterne imellem.

Det er ofte dyrt at udvikle sit eget system. Det er da også kun i forbindelse med store og langvarige projekter, at det overhovedet giver økonomisk mening at udvikle et redigeringsystem til en bestemt bog. I et forlag med et stort ordbogsprogram ville

det være helt uholdbart at udvikle og vedligeholde et redigeringsystem til hver ordbog, selv om de måtte være nok så forskellige.

Har man udviklet sit eget system, er der ikke andre kunder, der kan medvirke til finansiering af vedligehold og videre udvikling af redigeringsystemet. Og selv om man ikke skal have udviklet nye funktioner, har man med jævne mellemrum brug for at få tilpasset systemet til nye versioner af operativsystem e.l. Man har ikke en ekstern leverandør at holde økonomisk ansvarlig for, at systemet faktisk bringes til at virke – slet ikke inden for den aftalte tid. Projektorganisationen bliver sårbar; fx er det svært at erstatte en programmeringskyndig medarbejder, der forlader sit arbejde midt i processen. Det kan også være svært at afsætte tid til at få skrevet brugbar dokumentation. I det hele taget er de fleste ordbogsproducerende institutioner sjældent gearede til at fungere som softwareudviklere.

#### 4. Krav og prioriteringer

I dette afsnit vil jeg gennemgå, hvilke funktioner der med fordel kan prioriteres højt i forskellige situationer og ved forskellige projekt- og organisationstyper.

##### *4.1. Når man arbejder med flere ordbøger*

Når der i en institution udarbejdes flere forskellige ordbøger, bliver det naturligvis vigtigt, at redigeringsystemet kan håndtere mere end én ordbog ad gangen. Man skal kunne skifte mellem projekter, så en medarbejder ikke af tekniske grunde forhindres i at være tilknyttet flere projekter.

Det bliver også vigtigt, at de enkelte ordbøger kan have deres egne ordbogsgrammatikker. Hvis redigeringsystemet skal kunne yde redaktionel støtte (dvs. præsentere brugeren for netop de relevante strukturelle valg i en given kontekst) og stå for den strukturelle kontrol, giver det ikke mening at formulere en fælles grammatik for to potentielt meget forskellige værker.

##### *4.2. Når man skriver nye ordbøger*

Når man skriver en helt ny ordbog, vil man have stor glæde af, at redigeringsystemet kan tilbyde prædefinerede skabeloner – Bo Svensén kalder dem for ‘artikelmallar’ (Svensén 2004:499). En skabelon er et udvalg af informationstyper – typisk i form af ‘elementer’ eller ‘felter’ – der kan være tomme eller helt eller delvis udfyldt med tekst. Man kan forestille sig, at man har skabeloner for ordbogsartikler for ord, der tilhører en bestemt ordklasse og altid skal have præsenteret et bøjningsparadigme på en særlig måde. Eller man kan forestille sig skabeloner for forskellige slags henvisningsstrukturer, der kan forekomme i forskellige sammenhænge i ordbogsartiklerne.

Ved nyskrivning er det særlig vigtigt, at redigeringsystemet kan gemme artikler, der er ufærdige ifølge ordbogens grammatik. Det kan dels gælde en enkelt lang

artikel, der udarbejdes over flere dage, dels også det helt normale tilfælde, hvor man tilføjer én enkelt informationstype ad gangen til alle eller mange artikler på én gang. Ofte starter man med at oprette artikler, der blot indeholder opslagsord og evt. ordklasse, morfologi, etymologi eller anden information, der adskiller de enkelte homografer. Disse ufærdige artikler skal naturligvis kunne gemmes, uden at systemet bliver ustabil.

#### *4.3. Når man reviderer eksisterende ordbøger*

Arbejder man med nye udgaver af eksisterende ordbøger, er det andre forhold, der gør sig gældende. I den proces bliver avanceret søgning og statistikfunktionalitet ofte meget vigtig. Kan man kombinere søgning i struktur og indhold, er man efter min mening godt hjulpet. Man bør i den forbindelse vurdere, om søgesprogets syntaks lever op til det kompromis mellem brugervenlighed på den ene side og styrke og præcision på den anden, der passer bedst til brugerne. Jo mere avancerede søgninger man ønsker sig, jo mere synes syntaksen at komme til at ligne matematiske udtryk.

Hurtigt artikelopslag er uundværligt. Man har hele tiden brug for at kunne finde og åbne en hvilken som helst ordbogsartikel – ofte blot for at se, hvordan den er bygget op eller for at kopiere et strukturfragment. Moderne brugeres tålmodighed er ikke stor, så et opslag skal ganske enkelt være lynhurtigt, hvis ikke de skal blive utilfredse.

Det skal være nemt at orientere sig og bevæge sig rundt i den enkelte ordbogsartikels mikrostruktur. Der skal være et klart og pædagogisk indtastningsbillede, og der skal være relevante funktioner til at hoppe mellem eller flytte rundt på afsnit, niveauer, felter, elementer og andre komponenter, som strukturen måtte være bygget op af.

Hvis den udgave af ordbogen, man arbejder med, er den første, der redigeres i systemet, er der ofte en del strukturfejl i materialet. Det kan være, at man har konverteret fra et typografisk format (fx et tekstbehandlings- eller satsformat). Så vil der normalt være strukturer, der enten er fortolket forkert eller slet ikke er genkendt. Det er ønskeligt, at systemet fungerer stabilt alligevel og kan gemme artikler, selv om de indeholder mange strukturelle fejl.

Fejl kan naturligvis undgås ved, at man udarbejder ordbogsgrammatikken, så den bliver meget tolerant over for alle mulige strukturkombinationer. En sådan overtolerant grammatik kan ikke bruges til at støtte redigeringsprocessen på en sådan måde, at man kun får tilbudt relevante informationstyper i en given kontekst. Derfor vil det være en meget nyttig egenskab ved redigeringsystemet, hvis man kan knytte flere forskellige ordbogsgrammatikker til ét og samme ordbogsværk. Det kan fx være en overtolerant (eller ‘deskriptiv’) grammatik, der tillader alle de strukturer, der blev genkendt i ordbogen under konverteringen, og en ideel (eller ‘normativ’) grammatik, der beskriver den mikrostruktur, man faktisk ønsker, at alle artikler skal



følge. Undervejs i processen vil man nok have brug for at validere mod grammatikker, der beskriver forskellige mellemstadier.

#### ***4.4. Når man hyppigt udgiver nye udgaver af en ordbog***

Udgiver man nye udgaver af en ordbog med forholdsvis korte intervaller – hvilket nok oftere er tilfældet med kommercielle ordbøger end med mere videnskabelige ordbøger – vil man for hver udgave prioritere, hvilke opgaver der skal løses netop denne gang. De fleste ressourcer går som regel til at indføre nyt indhold og revidere eksisterende indhold i ordbogen. Derfor vil nogle af de strukturelle opgaver ofte løses over en række udgaver. Man kan sige, at ordbogen sjældent gøres helt færdig, før den udgives.

Det kan til mange formål være nyttigt at registrere, hvornår et givent ord kommer med i en ordbog. Også i forbindelse med markedsføring har dette værdi. Til dette formål vil man have glæde af en form for registreringsfunktion i redigeringsystemet.

#### ***4.5. Når flere forfattere arbejder på den samme ordbog***

Hvis flere skal skrive i den samme ordbogsbase, bliver de klassiske databasefunktioner vigtige: Systemet skal sikre, at der tildeles udelelig adgang til artiklerne. Og man skal kunne tildele forskellige brugere forskellige rettigheder. Nogle brugere skal have læse- og skriveret, nogle skal kun have læseret, og nogle værker må måske kun ses af udvalgte brugere.

Selv om man som forfatter kun arbejder med en del af bogens artikler (fx alle adjektiver, alle juridiske termer eller alle opslagsord, der starter med 'k'), har man også her behov for at kunne se alle artikler, når man har brug for det. Bl.a. skal man jo kunne lave krydshenvisninger på sikker vis. Man kan også forestille sig, at de udvalg af ordbogsartikler, som er tildelt de enkelte forfattere, overlapper hinanden.

Til denne form for projektorganisation bliver det nødvendigt med en veludbygget workflow-håndtering. Man skal kunne følge projektets fremadskriden, og man skal kunne se, hvem der har skrevet og rettet hvor. Man kan også forestille sig, at man har brug for at kende omfanget af den enkelte forfatters artikelproduktion for at kunne udbetale ham eller hende det korrekte honorar. Det siger sig selv, at administration og bogholderi, der føres i hånden – dvs. uden for systemet – bliver både mindre præcist og mere tidskrævende.

#### ***4.6. Brugerkompetence***

Der er stor forskel på, hvilke kompetencer de forskellige medarbejdere på en ordbog har. Nogle har store it-kundskaber, andre næsten ingen. Det kan skyldes alder, baggrund eller helt andre forhold. Den vigtigste kompetence hos en ordbogsforfatter er så afgjort den leksikografisk-sproglige kompetence.

Der er ligeledes forskel på, hvordan forskellige personer opfatter en ordbogsarti-

kel: Ser de den som et hierarki af indlejrede informationstyper, ser de den som en lineær sekvens af informationstyper, eller tænker de ganske enkelt på en ordbogsartikel som et stykke (typograferet) tekst? I figur 1 ses et strukturelt eksempel på alle tre opfattelser af en ordbogsartikel fra en tænkt dansk-fransk ordbog.

Hierarkisk præsentation	Sekventiel præsentation	Tekstpræsentation
<artikel> <hoved> <opslagsord> <b>ordforråd</b> <klasse> <b>sb.</b> <krop> <oversættelse> <b>vocabulaire</b>	opslagsord: <b>ordforråd</b> klasse: <b>sb.</b> oversættelse: <b>vocabulaire</b>	<b>ordforråd</b> <i>sb.</i> <b>vocabulaire.</b>

Figur 1.

At forfatterne har forskellige opfattelser af en ordbogsartikels natur, betyder ikke, at de ikke kan forstå og følge de samme redaktionelle retningslinjer. Jeg synes, det er vigtigt, at man i redigeringsystemet kan komme de enkelte forfattere så meget i møde, at de hver især stimuleres og støttes i deres artikelproduktion. Hvis redigeringsystemet tillader fleksibel præsentation af den samme artikel, kan man undgå at skulle udpege én konceptuel opfattelse som den ‘rigtige’. Det vil være meget ærgerligt at måtte sige nej tak til en knaldgod forfatter, fordi redigeringsystemet er for svært at bruge for ham eller hende.

#### 4.7. Når én ordbog publiceres på flere måder

At kunne vise ordbogsdata på flere måder kan også være en fordel, hvis man vil publicere sin ordbog på flere forskellige måder. Ofte vil man have flere faste linjebrud i en elektronisk ordbog end i en trykt bog. I den situation er der behov for at kunne se flere forskellige previews af ordbogsteksten. Det vil også være fint, om man kan definere, hvilke output-formater man vil have data eksporteret i. Det kan fx være html, xsl:FO, pdf eller et proprietært satsformat.

#### 4.8. Når man udveksler ordbogsdata med andre

Hvis man gerne vil kunne udveksle data med andre, skal kravet om eksport af data i et standardiseret dataformat prioriteres. Med til en ordbog hører – ud over selve data – ordbogsgrammatikken, alfabetiseringsnøglen og de forskellige outputdefinitioner. Hvis alle disse definitioner også kan eksporteres i et standardformat, kan man udveksle hele pakken. Andre, der kan håndtere samme format, kan så indlæse data og definitioner meget hurtigt i deres eget system og komme i gang med redigeringsarbejdet. Et oplagt og meget udbredt dataformat er XML<sup>5</sup> med de tilhørende

<sup>5</sup> eXtensible Markup Language

standarder for datas grammatik, DTD<sup>6</sup> og xml:schema, og standarden for XML-transformation XSLT.<sup>7</sup>

## 5. Konklusion

Det er sædvanligvis en stor investering at anskaffe sig et nyt redigeringsystem. Det kan derfor betale sig at bruge tid på at kortlægge sine behov og forventninger for at kunne vælge den løsning, der passer en bedst. Man har stor fordel af at formulere sine krav og ønsker og at opstille dem i prioriteret orden, så godt det nu kan lade sig gøre.

For mig at se er fleksibilitet en meget vigtig parameter. De vilkår, man arbejder under, har det nemlig med at ændre sig. Derfor er det meget vigtigt, at man kan ændre både ordbogsgrammatik og præsentations- og publiceringsdefinitioner undervejs. Da selve redigeringsprocessen uden sammenligning udgør den længste periode i et ordbogsprojekt, synes jeg, man skal prioritere redigeringsfaciliteterne meget højt. Det er disse faciliteter, der skal bruges i det daglige arbejde – dag ud og dag ind. Det er afgørende for datakvaliteten, at man indarbejder sine redaktionelle principper så godt som muligt i en ordbogsgrammatik for at kunne yde forfatterne mest mulig redaktionel støtte og samtidig holde en stram kontrol med datastrukturen.

I mine øjne er det vigtigt, at de valg, der kræver leksikografisk, redaktionel, pædagogisk eller æstetisk-grafisk kompetence, faktisk foretages af medarbejderne. Det kan være både svært og ærgerligt at skulle indrette sine ordbøger efter et bestemt stykke softwares (tilfældige) udformning. De vigtigste opgaver, der forudsætter sådanne valg, er:

- udformning af ordbogsgrammatik i overensstemmelse med de vedtagne redaktionelle retningslinjer
- fastlæggelse af principper for alfabetisering
- design af skærmpresentation til indtastning og preview
- fastlæggelse af typografi til forskellige udgivelser af ordbogsdata (fx som trykt bog, online-ordbog eller ordbog på mobiltelefon)

Redigeringsystemet skal tage sig af de mere tekniske, entydige og gentagne operationer og ikke mindst kontrol af data ifølge de regler, man har fastlagt.

Det er min påstand, at det er ulige meget lettere for en leksikografisk skolet person at lære sig en formel syntaks til beskrivelse af ordbogsdata, end det er for en programmør at forstå og træffe egentlige leksikografiske valg.

---

<sup>6</sup> Document Type Definition

<sup>7</sup> eXtensible Stylesheet Language Transformations

## Litteratur

### Anvendt litteratur:

- DWS 2004: *Third International Workshop on DICTIONARY WRITING SYSTEMS*  
<http://nlp.fi.muni.cz/dws2004/>.
- Ridings, Daniel 2003: Lexicographic workbench: A case history. I: Sterkenburg, Piet van (red.): *A Practical Guide to Lexicography*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 204-214.
- Svensén, Bo 2004: *Handbok i lexikografi*, 2. udgave. Stockholm: Norstedts Akademiska Förlag.

### Anden litteratur:

- van den Broek, Thijs 2005: *Choosing an XML editor* <http://ahds.ac.uk/creating/information-papers/xml-editors/>.
- Brooks, Kenneth P. 1988: *A Two-view Document Editor with User-definable Document Structure*. Palo Alto (California): Digital Systems Research Center.
- Harold, Elliotte Rusty 2002: *XML in a Nutshell – A Desktop Quick Reference*, 2. udgave. Cambridge (Massachusetts): O'Reilly.
- Joffe, David og De Schryver, Gilles-Maurice 2004: TshwaneLex, a state-of-the-art dictionary compilation program. I: Williams, Geoffrey and Vessier, Sandra (red.): *Proceedings of the Eleventh EURALEX International Congress*. Université de Bretagne Sud, Lorient, 99-104.
- The DANLEX-GROUP (Hjorth, Ebba/Nistrup, Bodil/Norling-Christensen, Ole/Jacobsen, Jane Rosenkilde/Ruus, Hanne) 1987: *Descriptive Tools for Electronic Processing of Dictionary Data*, Lexicographica Series maior 20. Tübingen: Max Niemeyer Verlag.