

Peter Widell og Mette Kunøe (udg.):
10. Møde om Udforskningen af Dansk Sprog
Århus 2005

MATERIALFILOLOGI OG XML-OPMÆRKNING HVEM MÆRKER OMKVÆDET?

Af Hanne Ruus (Københavns Universitet)

Materialfilologi

I de sidste snart 15 år har filologien fået en renæssance i form af materialfilologi.¹ Materialfilologien blev lanceret i 1990 med et særnummer af tidsskriftet *Speculum* hvis tema var New Philology, redigeret af den romanske middelalderfilolog Stephen Nichols (Nichols 1990), senere indførte Nichols termen Material Philology (Nichols 1997). Denne anden term er mere hensigtsmæssig da den ikke fører til forveksling med modstillingen af den klassiske filologi der behandler græske og latinske oldtidstekster, og Neuphilologie eller nyfilologien der arbejder med udgivelse af tekster på de europæiske folkesprog fra middelalderen og fremefter. På dansk foreslås termen materialfilologi for den teoretiske og praktiske nyorientering mod de fysiske tekstbærere og deres liv i samtid og eftertid. Den materialfilologiske tilgang er ikke kun relevant for håndskrevne tekster, også trykte tekster og deres opståen og reception i samtid og eftertid får fornyet aktualitet og interesse ved at blive studeret med materialfilologisk værktøj; D.C. Greetham påpeger i sin håndbog i *Textual Scholarship* hvordan de mange deldiscipliner der beskæftiger sig med bøger og tekster, kan befrugte hinanden ved at samarbejde (Greetham 1994:4-5).

Den materialfilologiske tilgang indebærer en ny fokus på de tekstbærende genstande, på manuskripter og bøger i deres helhed, deres fysiske fremtrædelsesform og dens betydning for teksterne i manuskript og bog, på skrivere, udgivere, samlere og trykkere og deres forhold til og interesse i de tekstbærere som de bidrog til fremstillingen af, på den enkelte tekstbærer og dens teksters relationer til andre førtidige og samtidige tekster og deres repræsentation, og endelig på manuskripters og bøgernes efterliv og reception i senere tider, bl.a. i form af udgaver og udgivelsesprincipper.

¹I Anne Mette Hansens ph.d.-afhandling: *Den danske bønnebogstradition i materialfilologisk belysning* (2004), kap 1., er det påvist at materialfilologien er en renæssance i ordets egentlige betydning idet en stor del af de synspunkter der nu fremføres, svarer ganske godt til J.N. Madvigs forestillinger om Filologiens Encyklopædi fra 1830'erne.

xml-mærkning

Den fornyede fokus på selve tekstens fremtrædelsesform og materielle omgivelser skærper interessen for detaljen og nødvendiggør en omhyggelig skelnen mellem hvad der faktisk står i teksten, og hvilke oplysninger udgiveren tilføjer. Her får man brug for markering af hvilken slags oplysninger der registreres. Efterhånden foreligger alle tekster i digital form, og man kan derfor benytte et standardiseret digitalt mærkningsprog; Standard Generalized Markup Language (SGML) blev udarbejdet allerede i 1980'erne, og det er i 2000-tallet næsten helt fortrængt af Extensible Markup Language (xml).

Ved at forsyne teksten i digital form med mærker efter xml-standarden eller en dermed kompatibel mærkning opnår man en række fordele:

- Ved at skelne mellem informationsbevarende og informationsberigende oplysningstyper får man en klar adskillelse af selve kildeteksten fra de oplysninger man tilføjer om den.
- Ved at specificere tekstens struktur i en dokumenttypedefinition (dtd) sikrer man at den enkelte tekst er fuldstændig entydigt beskrevet. I selve teksten anføres en specifikation af hvilke tegn der indgår i teksten. Dette er særlig påkrævet så snart man bevæger sig uden for repertoiret af tegn mellem a-z, A-Z, 0-9 og er dermed aktuelt for enhver dansk tekst.²
- Ved specifikationen af tekststruktur og tegnrepertoire bliver den digitale tekst uafhængig af det program den er frembragt med, og af den type pc den er oprettet på. Dermed sikrer man at det store arbejde som en præcis tekstgengivelse repræsenterer, også kan nyttiggøres i fremtiden.

Her følger til illustration et lille uddrag af en tekst mærket med xml-mærker. Teksten indgår i DUDS³-projektet hvor Dorthe Duncker og Hanne Ruus arbejder på at opbygge en digital tekstsamling med flerniveaumærkning – foreløbig kildetekstniveau og leksikalsk niveau – dækkende perioden ca. 1400 til nutiden. Først en version af tekstuddraget hvor man ser de bibliografiske oplysninger og de første tekstlinjer, mens mærkerne og det tilføjede leksikalske niveau er skjult, derefter de samme data med alle mærker vist.

Tekstart: A295

² Eksempler på tegn uden for a-z, A-Z, 0-9 og deres behandling findes i kap. 3 i Anne Mette Hansen 2004.

³ DUDS står for digitale undersøgelser af dansk sprog.

Titel: B.T.

Dato: Tirsdag 14. april 1970

Findested: s.5, art.1, spalte 1

Afslører russisk atom-ubåds // forlis: Så

den i vanskelighed ved Spanien

<tekstart>A295</tekstart><hoved><titel>B.T.</titel><dato><dag>Tirsdag</dag><datonummer>14.</datonummer><maaned>april</maaned><aar>1970</aar></dato><findested><side>5</side><artikel>1</artikel><spalte>1</spalte></findested></hoved>
<tekstuddrag><overskrift><periode><refer lbnr="1"><kildeord neutral="afslører">Afslører</kildeord><lemma metode="mltdic" form="afsløre" klasse="vb"/></refer><refer lbnr="2"><kildeord neutral="russisk">russisk</kildeord><lemma metode="mltdic" form="russisk" klasse="adj"/></refer><refer lbnr="3" ok="ja"><kildeord neutral="atomubåds">atom-ubåds</kildeord><lemma metode="morfana" form="atomubåd" klasse="sb"/></refer><start/><refer lbnr="4" ok="ja"><kildeord neutral="forlis">forlis</kildeord><lemma metode="person" form="forlis" klasse="sb"/><postsep>:</postsep></refer></periode><periode><refer lbnr="5" ok="ja"><kildeord neutral="så">Så</kildeord><lemma metode="homhjul" form="se" klasse="vb"/></refer><nl/><refer lbnr="6" ok="ja"><kildeord neutral="den">den</kildeord><lemma metode="mltdic" form="den" klasse="pron"/></refer><refer lbnr="7" ok="ja"><kildeord neutral="i">i</kildeord><lemma metode="homhjul" form="i" klasse="præp"/></refer><refer lbnr="8"><kildeord neutral="vanskelighed">vanskelighed</kildeord><lemma metode="ro86" form="vanskelighed" klasse="sb"/></refer><refer lbnr="9" ok="ja"><kildeord neutral="ved">ved</kildeord><lemma metode="homhjul" form="ved" klasse="præp"/></refer><refer lbnr="10"><kildeord neutral="Spanien">Spanien</kildeord><lemma metode="mltdic" form="Spanien" klasse="stn"/></refer></periode></overskrift>
... </tekstuddrag>

Man ser at de bibliografiske oplysninger holdes adskilt fra teksten ved hjælp af mærkeparret <hoved>...</hoved>, at tekstdelen er mærket som overskrift, og at selve kildetekstens ord optræder omgivet af <kildeord>...</kildeord> ledsaget af attributten neutral med ordets ortografisk neutrale form, f.eks. <kildeord neutral="atomubåds">atom-ubåds</kildeord>. I det følgende vil jeg med en folkevisersion fra *Svanings håndskrift I* som eksempel demonstrere hvordan materialfilologiske synspunkter og tekstmærkning kan bringes til spille sammen med nye indsigter til følge.

En folkeviser som eksempel

I forskningsprojektet *Dansk Folkevisekultur 1550-1700*, hvis resultater er fremlagt i værket *Svøbt i mår* 1-4 (Lundgreen-Nielsen og Ruus 1999-2002), var den ældste viseoverlevering før den første trykte visesamling fra 1591 genstand for indgående undersøgelser, både ud fra materialfilologiske synspunkter (Ruus 2003) og med anvendelse af tekstmærkning (if.

SGML-standarden da xml først kom frem et stykke ind i projektperioden 1993-1998). Alle tekster fra de ældste visebøger blev behandlet og er udgivet på den cd-rom der følger med bind 3 af *Svøbt i mår*. I det følgende bruges visen "Jomfruen og Dværgekongen" fra *Svanings håndskrift I*, nr. 93, side 161^r trykt i *Danmarks gamle Folkeviser* som nr. 37, som tekst Aa til at eksemplificere samspillet mellem materialfilologiske synspunkter, digital (re)præsentation og tekstmærkning.

Cd-rom-udgivelsen tilgodeser en række materialfilologiske synspunkter:

- Man kan studere de tekstlige omgivelser i manuskriptet. I katalogerne på *Svøbt i mår*-cd-rommen findes oplysninger om alle tekster i alle de behandlede visebøger og deres placering i visebøgerne tilgængelige og søgbare. F.eks. kan man bemærke at "Jomfruen og Dværgekongen" i *Svanings håndskrift I* godt nok er omgivet af ridderviser, men blandt de ti foregående viser findes en episk vise, og blandt de ti efterfølgende findes tre episke viser og to lyriske.
- Man kan studere selve manuskriptets fremtrædelseform idet alle viser der ikke er trykt i de videnskabelige udgaver, er læst efter billeder af håndskriftet, som er lagret i digital form på cd-rommen.
- Man får oplysninger om de enkelte visers reception i samtiden i det 16. og 17. århundrede idet der er linket mellem alle versioner af en vise i de visebøger der er registreret i katalogerne. F.eks. kan man se at "Jomfruen og Dværgekongen" findes i to andre visebøger fra det 16. århundrede og i to visebøger fra begyndelsen af det 17.
- Der er præcise henvisninger til de enkelte visers reception i eftertiden dels ved henvisninger til deres behandling og placering i de videnskabelige udgaver, dels ved et viseregister i *Svøbt i mår* 4 der giver henvisninger til de steder i værkets fire bind hvor de enkelte viser er omtalt.

På *Svøbt i mår*-cd-rommen er samtlige viser i den ældste overlevering, 1014 viseversioner, mærket på ordniveau efter SGML-standard, og over halvdelen er forsynet med et leksikalsk niveau indeholdende ortografisk neutral form, lemma og ordklasse til hver ordform (Duncker og Ruus 2000, Ruus 2001). På basis af kildenniveau og leksikalsk niveau er der opbygget en komplet ordbog over alle ord i viserne, som forskere og studerende med vidt forskellige faglige interesser kan bruge til at undersøge viser og viseudtryk på kryds og tværs (Ruus 2002). Et par eksempler fra "Jomfruen og Dværgekongen" fra *Svanings håndskrift I* kan demonstrere den anvendte mærknings potentialer:

På basis af de ortografisk neutrale ordformer kan der fremstilles en nutidslæservenlig

version af teksten. Her følger to udsnit af visen i denne form. I visen er jomfruen blevet forført af en dværg og lever med ham om natten i bjerget og med sin moder om dagen.

33.

“hør du, kære datter min:
og hvi da falmer dine fagere skind?

34.

hvi da falmer din kinder så?
hvi kan dig ingen ridder få?

35.

dig haver bedet konger og grever rige:
hvi lader du ingen være din lige?”

36.

“min kære moder, vredes ikke ved:
alle min sorrig siger jeg eder.

37.

det duer ikke at dølge for eder:
den dværrigkonge haver lokket mig.

38.

jeg haver mig i bjerget inde
syv sønner ved den dværrig så fin.

Dværgen hører samtalen og tvinger jomfruen tilbage til bjerget. Efter afsked med forældre, venner og frænder vender jomfruen tilbage til bjerget:

52.

den jomfru ind i bjerget tren:
hendes syv sønner går hender igen.

53.

den ungeste går hender grædendes imod,
den ældste sætte hende hynde og stol.

54.

“hør I, kære moder min:
hvi da røbte I fader min?”

55.

så såre da randt hender tåre på kind:
hun gade ikke svaret sønner sine.

56.

hun gad dem ikke svaret for sorrig og nød:
der tiden gøres mig så lang.

og end før midnat var hun død.
eja, sorrigen, hvad du est tung!

Viseteksterne på cd-rommen bygger på de videnskabelige udgaver og anfører derfor kun mellemkvæd og omkvæd ved første og sidste strofe. Når man på god materialfilologisk vis kontrollerer viseteksten i Svanings Håndskrift s. 161^{r følgende}, viser det sig at mellemkvædet er anført i alle strofer. Omkvædet anføres – også i selve visebøgerne – ofte kun ved få strofer og antydes ellers med første bogstav af ordene i omkvædet ved de øvrige strofer. Fra en materialfilologisk synsvinkel bør en visetekst altså fremlægges med mellemkvæd og omkvæd anført ved alle strofer, kun på denne måde får man adgang til visen som den fungerer når den synges eller læses op. Når teksten er forsynet med mærker, kan man tilføje de manglende mellemkvæd og omkvæd ledsaget af en oplysning om at de er indsat i forhold til den trykte udgaves tekst; i eksemplerne i det følgende kursiveres den indsatte tekst. Her kommer to strofer med indsat mellemkvæd og omkvæd. Det er dværgen der beder sin moder om råd til at lokke jomfruen.

14.
Duerigen suøber sitt hoffuit y skiend,
Der tiden giøris mig saa lang.
hand ganger y lofft for sin moder ind.
Eya, huad sorigen du est tung!

15.
"Min kiere moder, y kiender mig rad:
Der tiden giøris mig saa lang.
huor ieg skall kongens datter foe.
Eya, huad sorigen du est tung!

Når mellemkvædet kommer ind i alle strofer, bliver det klart at denne følelse er vigtig for visen. Det kan da være oplysende at undersøge om formuleringen af at tiden går langsomt, forekommer i andre viser udtrykt med netop disse ord. En søgning på *Svøbt i mår*-cd-rommen efter ordene tiden-gøre-mig-lang i samme verslinje og vilkårlig rækkefølge giver elleve eksempler fra ni viseversioner – de syv fra lyriske viser – og afslører at det som oftest er længsel eller sorg der er anledning til følelsen, f.eks. "af længelse gøres tiden mig lang" eller "hemmelig længsel, hvad du est tung, | hvad du est tung at bære; | du gør mig tiden og stunden så lang, | for hvem skal jeg mig kære;"; dette passer jo smukt med sorgen i folkevisens omkvæd: "Eya, huad sorigen du est tung!".

En læsning af visen med de indsatte mellemkvæd og omkvæd giver inspiration til at se nærmere på de mange gentagelser og deres variationer. Ved mellemkvædet er det oplagt at variationen ligger i hvem tiden gøres lang for; mærkningen suppleres derfor med en attribut til

det indsatte mellemkvæd der hedder føler_mellemkvæd med visens personer som værdier. Andre gentagelser udmærker sig ved en variation i ordlyd; for at registrere dette med mærker indføjes et gentagelseselement med attributter. Her følger den del af visens xml-dtd der viser de indføjede oplysninger om omkvæd, mellemkvæd og gentagelser:

```
<!ELEMENT indsat EMPTY>
<!ATTLIST indsat
    mellemkvæd CDATA #IMPLIED
    føler_mkvæd CDATA #IMPLIED
    omkvæd CDATA #IMPLIED>
<!ELEMENT gentagelse ( #PCDATA )*>
<!ATTLIST gentagelse
    delnummer CDATA #REQUIRED>
<!ELEMENT gentagelsesvariation ( #PCDATA | betydning | formel )*>
<!ATTLIST gentagelsesvariation
    delnummer CDATA #REQUIRED>
```

Ved en digital fremvisning af den mærkede tekst anvendes forskellige farver til at vise de forskellige attributværdier. I de følgende eksempler er følgende notationskonvention brugt: indsat tekst (mellemkvæd og omkvæd) *kursiveres*; jomfruen føler i **fed stor skrift**, jomfruens moder føler i **fed normal skrift**, og jomfruens yngste søn føler i **fed lille skrift**. Gentagelser med variation i ordlyden en skrevet med skriften **Serifa BT i fed**, de udtryk der varierer, er gjort **store**, og anden forekomst er understreget. De fleste af de følgende strofer er anført i eksemplet på en nutidslæservenlig udgave af visen ovenfor, en sammenligning viser hvor meget man går glip af uden de indsatte mellemkvæd og omkvæd.

30.

Hinder beder rige konger fem,
Der tiden giøris mig saa lang.
oc ingen vill hun haffue aff dem.
Eya, huad sorigen du est tung!

31.

Hinder beder ryge greffuer ny,
Der tiden giøris mig saa lang.
oc ingen vill hun haffue aff dy.
Eya, huad sorigen du est tung!

32.

Iomfru Ermerig och hindis moder

Der tiden giøris mig saa lang.

saa ene syder di y dieris bure.

Eya, huad sorigen du est tung!

33.

“Hør du, kiere datter min:

Der tiden giøris mig saa lang.

och hui daa falffmer dine faffuere skind?

Eya, huad sorigen du est tung!

34.

Hui daa falmer din kiender saa?

Der tiden giøris mig saa lang.

hui kand dig ingen ryder foe?

Eya, huad sorigen du est tung!

35.

Dig haffuer beditt konger och greffuer rige:

Der tiden giøris mig saa lang.

hui lader du ingen vere din lige?”

Eya, huad sorigen du est tung!

36.

“Min kiere moder, vredis ichi vid:

Der tiden giøris mig saa lang.

alle min sorig siger ieg eder.

Eya, huad sorigen du est tung!

37.

Ditt duer iche at dulle for eder:

Der tiden giøris mig saa lang.

den duerig-konge haffuer lochit mig.

Eya, huad sorigen du est tung!

[...]

45.

Hun klapitt paa loffuis[-]dør met sin skind:

Der tiden giøris mig saa lang.

“Gode[-]natt, kiere fader och moder myn!

Eya, huad sorigen du est tung!

46.

Gode[-]natt, fader oc moder min!

Der tiden giøris mig saa lang.

min søster oc min broder!

Eya sorigen huad du est tung!

47.

Och der-till venner flere,

Der tiden giøris mig saa lang.

och venner och frender till[-]lige.

Eya, huad sorigen du est tung!

48.

Och segne eder **Christ hin rige:**

Der tiden giøris mig saa lang.

y syer mig aldrig miere!

Eya, huad sorigen du est tung!

49.

Och segne eder **alt dett, Gud haffuer skabt!**

Der tiden giøris mig saa lang.

och beder løche saa vor mig act. ”

Eya, huad sorigen du est tung!

[...]

52.

Den iomfru ind y biergitt trend:

Der tiden giøris mig saa lang.

hindis VII sønner gaar hinder igen.

Eya, huad sorigen du est tung!

53.

Den ungiste gaaer hinder gredendis imod,

Der tiden giøris mig saa lang.

den elste sette hinde hønede och stoell.

Eya, huad sorigen du est tung!

54.

“Hør y, kiere moder min:

Der tiden giøris mig saa lang.

hui da røbte y fader min?”

Eya, huad sorigen du est tung!

55.

Saa sore daa rand hinder tore paa kiend:

Der tiden giøris mig saa lang.

hun gade ichi suarit sønner sine.

Eya, huad sorigen du est tung!

Konklusion

Den materialfilologiske tilgang til undersøgelser af en enkelt viseversion og mulighederne for at supplere en xml-mærket tekst med oplysninger inspireret af materialfilologien har demonstreret at filologiens nyorientering kan få stort udbytte af xml-mærkningsteknologien fordi den muliggør fastholdelse af hermeneutiske indsigter på de relevante steder i teksten, klart adskilt fra selve kildeordlyden. Undersøgelserne har også vist styrken i at arbejde med omhyggeligt mærkede tekster – det lille kig ud i den større visesamling gav større klarhed over mellemkvædets betydning. Endelig har kombinationen af materialfilologiske synspunkter og xml-mærkning inspireret til nye iagttagelser i teksten som vil kunne frugtbares i andre tekster; den præcise registrering af gentagelser og deres variation ikke kun i ordlyd, men også i kraft af deres placering i visens kontekst, ansporer til tilsvarende undersøgelser af andre viser. Kombinationen af materialfilologi og xml-mærkning har et klart heuristisk potentiale.

Litteratur

- Duncker, Dorte and Hanne Ruus (2000) Multi Level Text Representation in an LCB, in *Symposium on Lexicography IX, Proceedings of the Ninth International Symposium on Lexicography April 23-25, 1998 at the University of Copenhagen*, Edited by Jens Erik Mogensen, Viggo Hjørnager Pedersen and Arne Zettersten, Lexicographica, Series Mayor, Band 103, Max Niemeyer Verlag, Tübingen, p. 77-97.
- Duncker, Dorte, Jesper Juul og Hanne Ruus (2001) *Svøbt i mår, Dansk Folkevisekultur 1550-1700, cd-rom*, i Flemming Lundgreen-Nielsen og Hanne Ruus (red.) (2001).
- Greetham, D.C. (1994) *Textual Scholarship, An Introduction*. New York & London: Garland Publishing Inc.
- Hansen, Anne Mette (2004) *Den danske bønnebogstradition i materialfilologisk belysning*, ph.d.-afhandling, Københavns Universitet.
- Lundgreen-Nielsen, Flemming og Hanne Ruus (red.) (1999) *Svøbt i mår*, bind 1: Adelskultur og visebøger, C.A. Reitzel, 428 s.
- Lundgreen-Nielsen, Flemming og Hanne Ruus (red.) (2000) *Svøbt i mår*, bind 2: Et spørgsmål om stil, C.A. Reitzel, 510 s.
- Lundgreen-Nielsen, Flemming og Hanne Ruus (red.) (2001) *Svøbt i mår, Dansk folkevisekultur 1550-1700*, bind 3: Tæt på viseteksterne, C.A. Reitzel, København, 516 s. med indlagt cd-rom.
- Lundgreen-Nielsen, Flemming og Hanne Ruus (red.) (2002) *Svøbt i mår, Dansk folkevisekultur 1550-1700*, bind 4: Lærdom og trolddom, C.A. Reitzel, København, 796 s.
- Nichols, Stephen G. (ed.) (1990) *Speculum: A Journal of Mediaeval Studies*, Vol 65 No. 1, Cambridge Massachusetts.
- Nichols, Stephen G. (1997) "Why material philology?" Some Thoughts by Stephen G. Nichols, Baltimore, Philologie als Textwissenschaft: Alte und neue Horizonte, *Zeitschrift für Deutsche Philologie*, 116, Sonderheft, Berlin: Erich Schmidt Verlag, s. 10-30.
- Ruus, Hanne (2001) Elektronisk viseforskning, i Flemming Lundgreen-Nielsen og Hanne Ruus (red.) (2001), s. 9-46.

- Ruus, Hanne (2002) A Corpus-based Electronic Dictionary for (Re)search, in Anna Braasch and Claus Povlsen (eds.): *Proceedings of theTenth EURALEX International Congress, EURALEX 2002*, Copenhagen, Denmark, August 13-17, 2002, Volume I, p.175-185.
- Ruus, Hanne (2003) The Oldest Danish Ballad Tradition in its Socio-cultural Setting, in *Text and Context: Combining Philology, History and Cultural Studies, Nordistica Tartuensis* 8, ed. by Kristel Zilmer and Stig Örjan Olsson, Tartu, Estonia, s. 121-131.