

Tekstlaboratoriet

Det historisk-filosofiske fakultet

Boks 1102 Blindern

0317 Oslo

Innhold

- Gjest i dette nummer
- Tekstlaben i Salzburg
- Seminar med HF og UF
- Oxford English Dictionary
- Tekstlab-ingeniøren informerer:

Tekstlaben på Internett

- Nye tilbud
- Kurs
- Kort om Tekstlaboratoriet

**Tekstlaboratoriet**

Det historisk-filosofiske
fakultet

ΒΥΛΛΕΤΙΝ
Nr. 3/november 1995 • 3. årgang

UNIVERSITETET I OSLO

• Gjest i dette nummer

Navn:	Egil Kraggerud
Stilling:	Professor i klassisk filologi
Sivilstatus:	Gift, tre sønner og en datter
Hobby:	Klassisk musikk



Gjesten denne gang er klassisk filolog i enhver forstand. Det klassiske dekker både gresk og latin, og det filologiske både språk og litteratur. De nymotens skillene som vi i dag opererer med mellom for eksempel språk og litteratur, er Egil Kraggerud glad han slipper å forholde seg til. Men datateknologien ser han som et ubetinget gode.

- Kan du si litt om hva du er opptatt av i forskningen?

- Jeg skrev doktoravhandling om dikteren Vergil, som levde på Augustus' tid like før år 0. Siden har jeg alltid vært interessert i den klassiske perioden, og har skrevet bøker og artikler om Vergil og Horats. Men jeg er også opptatt av middelalderen og alt som ble skrevet på latin da, kirkefedrene og den middelalderlatinske diktningen ikke minst.

Et av hovedproblemene for oss som driver med eldre tekster, er at vi aldri helt vet hva som er den "riktige teksten". Før boktrykkerkunsten ble jo alle tekster overlevert som avskrifter av avskrifter gjennom ofte mer enn 1000 år. Enhver skriver tolket forrige skrift på sin måte, og gjorde vel gjerne noen feil også. Slik har vi svært mange utgaver av hver tekst, og bortsett fra de som ble skrevet inn på stein, er det et puslespill å lete seg bakover. Det er dessuten varierende hvor mye som

er bevart, noe er helt ødelagt.

Når man i dag trykker tekstene i vitenskapelige utgaver, er det på hver side et stort fotnoteapparat som viser til andre tekster og tolkninger av ord og fraser. (Kraggerud henter straks et par utgaver ned fra hyllen.) For hver nye vitenskapelige utgave er man forhåpentligvis kommet litt nærmere den opprinnelige teksten. Jeg kan ta et eksempel fra noe jeg har arbeidet med nå i høst.

I Messias-eklogen, som er et dikt av Vergil, forekommer denne frasen i den siste vitenskapelige utgaven: "cui risere parentes", som vi i denne sammenhengen må oversette med: "foreldrene må smile til barnet". Imidlertid har jeg selv studert nøye et indirekte tekstvitne, og funnet at frasen faktisk må være nesten det motsatte: "qui risere parenti", altså at det er barnet som smiler til moren. Hvis det er slik jeg tror, må teksten rettes mot hele overleveringen. Det ville være svært verdifullt å få samlet alt tekstkritisk materiale til tekstene på CD-rom.

- Kan elektroniske verktøy virkelig være til nytte for en klassisk filolog?

- Her må jeg svare både et tja og et klart ja. Når det gjelder forskningen i den klassiske perioden, er det faktisk ikke fullt så nødvendig. Her har vi massevis av hjelpemidler. I tillegg til de vitenskapelige utgavene har vi konkordanser av mange slag. (Kraggerud henter en bibel-konkordanse og en Vergil-konkordanse ned fra hyllen.) Her står alle ordene i teksten under hverandre med kontekst og stedshenvisninger.

For middelalderen, derimot, er det ikke utgitt slike hjelpemidler. Men Det historisk-filosofiske fakultet og Det teologiske fakultet har nettopp gått sammen om å kjøpe en nylig utgitt elektronisk utgave (CD-rom-plater) av den såkalte Migne-patologien (opprinnelig utgitt på 1800-tallet), som inneholder en stor del av de europeiske latintekstene fra senantikken og middelalderen. Dette edb-materialet vil hjelpe oss kolossalt.

Selv er jeg faktisk en nykommer på edb-området. Jeg fikk PC først i 1989. (Kraggerud smiler litt.) Jeg tilhører de trege. Men jeg innser at jeg trenger å bruke de nye hjelpemidlene.

- Fins det noen norske latintekster?

- Åja. Det fins en del historiske tekster om konger, om Olavsdyrkelsen, og helgendyrkelsen i sin alminnelighet. Det dreier seg om ca. 200 sider tekst. (Kraggerud henter en skriftlig utgave ned fra hyllen.) Dessverre var reformasjonen på 1500-tallet mer systematisk her enn i andre land, svært mye av det kirkelige og religiøse ble ødelagt.

Det norske materialet er spennende, og vi vet lite om det. Vi vet for eksempel ikke nok om hvilke verker forfatterne hadde lest, og dermed hvem de hadde kontakt med. Når vi får benyttet den elektroniske Migne-patrologien, kan vi slå opp og søke der, og finne ut ting på en-to-tre som det ellers ville tatt mange år å undersøke. Til nå har forskningen vært usystematisk fordi det ikke har vært nok tilgjengelige hjelpemidler.

La meg ta et eksempel. I teksten til munken Theodricus (senere den samme som erkebiskop Tore, tror vi) fra 1100-tallet finner vi uttrykket "matrix abyssus" (moderdyppet). Dette var et underlig uttrykk, og jeg visste ikke hvor det kom fra, til jeg ved en tilfeldighet fant at det samme uttrykk var brukt av en forfatter i Paris, og at det hadde betydningen "det dyp som frembringer (alt vann)". Dette er med og bekrefter at det var kulturelle forbindelser mellom Norge og Paris på den tiden.

Som en følge av det nye elektroniske materialet, skal jeg nå inngå et forskningssamarbeid med tre andre fra Oslo og Bergen, hvor målet er å utgi en vitenskapelig utgave av de norske tekstene. To skal arbeide heltid og to deltid i fire år.

- Hva med fremtiden?

- Vi har ikke så rent få studenter. På semesteremnet i latin er det 40-50

studenter, på middelalderkurset er det 15-20, og på grunn- og mellomfag 8-10. I tillegg er det til enhver tid 1-2 hovedfagsstudenter og nå en doktorgradsstudent.

Rent generelt tror jeg at edb-materialet er nødvendig for å holde på nye studenter. De synes de nye perspektivene er spennende. Terskelen for å gå i gang med forskning blir lavere. Og det er en forutsetning for at vi skal kunne tilby en treårig dr.art.-grad. Jeg ser det elektroniske materialet som et utrolig viktig instrument.



• Tekstlaben i Salzburg

I september deltok Tekstlaboratoriets forsker på en internasjonal konferanse i Salzburg om gresk lingvistikk. Foredraget som ble presentert, hadde som utgangspunkt forskning som var gjort med Tekstlabens ressurser. Imidlertid hadde det også en teoretisk lingvistisk side, og denne kombinasjonen gjorde at foredraget ble tatt vel imot.

Rent konkret var problemstillingen subjektets plass i en spesiell type underordnede setninger, såkalt konjunktive setninger. Det har i faglitteraturen vært fremhevet at slike setninger ofte har et subjekt som kommer etter verbet. Noen ganger har det også vært hevdet at dette er eneste mulighet. Nettopp en slik uenighet om data er jo noe som egner seg ypperlig til å bruke korpus på, for å finne nærmere ut av den faktiske bruken. Det forskeren fant, var svært interessant: I de nevnte setningene hadde over 80 prosent et etterstilt subjekt, i direkte motsetning til hovedsetninger og andre typer underordnede setninger, som hadde over 80 prosent foranstilt subjekt. Dette betyr jo at det er en svært klar tendens å gjøre rede for, samtidig som man også må kunne gjøre rede for de resterende i underkant av 20 prosentene i begge typer setninger.

Den analysen som ble foreslått, bygget delvis på en syntaktisk analyse av GB-typen, og delvis på en mer pragmatisk orientert analyse med begrepet "fokus" som sentralt. I ettertid blir det et samarbeid mellom oss, Sverige og Spania som et resultat og en oppfølging av denne forskningen. Her må det understrekes at nettopp det grundige datagrunnlaget (korpuset besto av en mengde avisartikler) gjorde et overbevisende inntrykk.

Salzburg er en nydelig by, som tok seg flott ut med høstfarger, en elv med klart, grønt vann som rant gjennom sentrum, en gammel middelalderborg litt oppe i åsen, og kakepyntdekorerte bygninger. Konferansen hadde flere hundre deltagere fra hele verden, men bortsett fra to svensker, var det bare Tekstlab-forskeren fra Norden.



Seminar med HF og UF

I slutten av september ble det arrangert et todagers seminar for redaktører, designere og andre representanter fra Universitetsforlaget og folk som driver med elektroniske medier i undervisning og forskning ved HF i Oslo. Målet var at Universitetsforlaget skulle få se hvilke ting som rører seg på universitetet, og hvor de eventuelt kan sette inn ressurser eller være med å samarbeide.

Tekstlaboratoriet ble presentert av forskeren. Ellers var det mange interessante foredrag: Christian-Emil Ore presenterte Dokumentasjonsprosjektet. Han fortalte om de mange rundt om i Norges land som har fått arbeid i prosjektet for å registrere håndskrevet og annet materiale på de forskjellige områder, f.eks. arkeologi og leksikografi. Nina Kristiansen viste hvordan Netscape-programmet kan brukes til å finne interessante ting på nettet såvel rundt om i verden som i Dokumentasjonsprosjektets interessante arkeologisider. Hallvard Dørum, KRI, fortalte om "det elektroniske klasserom", og hvordan en

lærer kan undervise samtidig på flere geografiske steder vha avanserte videokameraer (som filmer den som tar ordet) og skjermer i begge klasserom. Sheila Coulson, Nordisk arkeologi, presenterte undervisningsprogrammet "Steinalder Studiesamlingen", et avansert interaktivt program med bilder, tekst og oppgaver.

Jan Oldervoll, Historisk institutt, UiB, fortalte om Absalonprosjektet, et imponerende undervisningsopplegg på nettet. Her kunne grunnfagsstudentene skrive oppgaver, få dem lest av lærere og hverandre, og ha faglige diskusjoner. Opplegget har ført til en dramatisk forbedring i strykprosenten og gjennomføringsprosenten. Jon Lanestedt, Prosjekt for nettbaserte hovedfagsstudier i nordisk språk og litteratur og i historie, fortalte hvordan dette prosjektet tenkes gjennomført i praksis. Signe Marie Sanne, Humanistisk Datasenter, Bergen, demonstrerte multimedieprogrammer for språk og kulturundervisning. Bård Sandvei fortalte om et fjernundervisningsprosjekt for semesteremne i oversettelse. Adm. dir. Trygve Ramberg foredro om hva et forlag som UF kan gjøre for å tilrettelegge for elektroniske læremidler for høyere utdanning.

To foredrag var mer polemiske. Dekanus, professor Bjarne Hodne snakket om konsekvensene for kulturfagene ved digitalisering. Han pekte på at det kan føre til fall i kompetanse, om man ved å miste nærheten til originalene mister f.eks. evnen til å lese håndskrift, eller arkivkunnskap. Videre kan det føre til svekket kildekritikk. På den annen side vil bedret kildetilgang føre forskningen mot det analytiske, man vil få en metodeutvikling, og ny kunnskap. Foredraget førte til en viss debatt, og Jan Oldervoll fra Bergen påpekte at man ved digitalisering tvert i mot kan komme *nærmere* kildene og arkivene enn før. Tidligere var det bare noen få som hadde tilgang til arkivene, og man måtte ofte nøye seg med avskrift. Digitalisering kan gjøre arkivene, med kataloger, mer tilgjengelige, og i tillegg til avskrift kan man få arkivmaterialet avbildet.

Professor Stig Johansson, IBA, snakket om elektroniske tekstkorpora. Hva kan de brukes til, hvordan kan de berike språkforskningen? Han la

stor vekt på at det idag er altfor vanskelig å få tillatelse til å lagre tekster elektronisk. Dette fører til at det for et lite språk som norsk blir svært vanskelig å få tak i brukbart materiale. Hvordan skal man kunne forske i norsk språk og litteratur med moderne metoder og hjelpemidler (altså slike som vi tilbyr på Tekstlaben) hvis det ikke finnes noe materiale å forske i? Her må forlagene komme på banen. Den svært strikte praksisen de fører i dag, må endres.

Oxford English Dictionary



Oxford English Dictionary (OED) er nå tilgjengelig i elektronisk form ved Tekstlaben. I oktober arrangerte Tekstlaboratoriet en innføring hvor Stig Johansson fra IBA ga et underholdende foredrag om bakgrunnen for OED, og det etterpå ble gitt en demonstrasjon i både PC- og Mac-utgavene.

Er man opptatt av definisjoner, etymologi og belagt bruk av ordene i det engelske språket, vil OED være til stor glede og nytte. For nordmenn kan det være interessant at Nansen er sitert ved hele fire anledninger. Ibsen er ikke sitert, men forekommer i og med ord som *ibsenesque* og *ibsenism*.. Ellers kan det nevnes at *Oslo* er kjent p.g.a. *the Oslo-breakfast*.

PC-brukere må kontakte Tekstlabens ingeniør for å få installert et par ting på maskinen sin før de kan ta i bruk OED via Windows. Mac-brukere kan ta den direkte i bruk ved å velge MacTekstlab og OED2 fra Velger-menyen. OED kan altså brukes over nettet.

Tekstlab-ingeniøren informerer:

Tekstlaben på Internett



Tekstlaben har hjemmeside på nettet, på World Wide Web (WWW),

også kjent som Internett. WWW er et nettverk av tjenester, hvor man med noen få tastetrykk eller museklikk kan navigere seg rundt i en overveldende mengde informasjon. Informasjonen presenteres sidevis. En side på nettet tar omlag en skjerm-side. Hovedsiden eller førstesiden til en organisasjon eller person kalles gjerne hjemmeside.

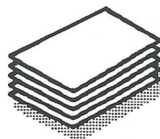
På sidene er det gjerne referanser ("pekere") til relevant informasjon. Referansene er markert som understrekinger i teksten, eller de kan være i form av bilder. Når man klikker med musa på en referanse, hentes den refererte siden inn, og blir presentert på skjermen. Adressen til Tekstlabens hjemmeside er:

<http://www.hf.uio.no/tekstlab/>

På sidene våre er det informasjon om oss, og om det vi tilbyr av tekster og programmer. Det kan være vanskelig å finne fram på nettet, så vi har laget en egen side med referanser til sider på nettet som kan være av interesse for språkfolk. Vi på Tekstlaben bruker referansesiden som utgangspunkt for å navigere på nettet, vi håper at andre har nytte av den også. På siden finnes referanser til viktige tekstsentre og tekstprosjekter. Referanser til sider for oppslag i bokmåls- og nynorskordboka, pluss Webster og Roget's Thesaurus finnes også på siden. I tillegg har siden referanser til sider for å søke etter sider og adresser på nettet: man oppgir et stikkord og får presentert en side med referanser til sider som inneholder stikkordet. Dette er nyttig hvis du vil søke etter noe selv, og fordi oversiktssider på nettet ikke alltid er ajourførte.

Netscape er det mest brukte programmet til å bevege seg på WWW. Netscape går på Macintosh og under Windows på PC. Hvis du trenger hjelp til Netscape, spør IT-ansvarlig på ditt institutt.

Helge Hauglin.



Nye tilbud

Tekstlaboratoriet har som en overordnet målsetning å skaffe til veie flere elektronisk tilgjengelige tekster på norsk. I motsetning til hva som er tilfelle for andre språk, er det jo ikke noe annet sted forskerne kan henvende seg når de trenger norskspråklig materiale. Vi er glade for å kunne fortelle at vi nå har en god del flere sakprosa-tekster. Det har vi fått tak i ved at Tekstlabens ingeniør har "lastet ned" dokumenter som har vært tilgjengelige over nettet, og som han siden har kjørt igjennom selvlagede programmer for at de skulle bli pene og ryddige. Det dreier seg om:

- 32 offentlige rapporter (NOU - Norges Offentlige Utredninger), 22 fra 1995, 10 fra 1994. Dette utgjør mye tekstmateriale, da rapportene gjennomsnittlig er en god del større enn f.eks. vanlige romaner.

Videre er vi i gang med å innhente tillatelse til å bruke en del nyere romanfragmenter fra norske forfattere. Det er den norske delen av det norsk-engelske parallellkorpuset til IBA som HFs ansatte og studenter etter hvert skal få tilgang til.

Vi har også meldt Tekstlaben inn i LDC (Lexical Data Consortium). Her vil vi etterhvert få både korpora (engelske og andre) og ordbøker gratis ved å betale en årlig avgift.



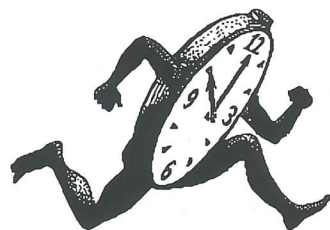
Kurs

Tekstlaben vil igjen holde et kurs i grunnleggende programvare for fakultetets ansatte og høyeregradsstudenter. Både tidligere kursdeltagere som føler behov for en oppfriskning, og nye deltagere, er hjertelig velkommen. Som vanlig blir det både demonstrasjon og praktiske

øvelser for hver enkelt. Det vil bli holdt i annen uke av desember: Mandag 4. og tirsdag 5. desember er det kurs for PC-brukere, og onsdag 5. og torsdag 6. desember er det kurs for Macintosh-brukere. Alle dager er det kl. 9.15-12.00. Kurset annonseres separat, så følg med.

I januar skal vi atter holde kurs for hovedfagsstudenter i forbindelse med hovedfagsstartkurset. Med tanke på dette er det kanskje også noen veiledere som føler behov for selv å gå på kurs. Husk da desember-kurset!

Kort om Tekstlaboratoriet



Tekstlabens fremtid

Tekstlaben har i dag to ansatte, en forsker (dr. philos.) og ingeniør (cand. scient.). Forskerstillingen er delt mellom kursvirksomhet ved og drift av Tekstlaben, og forskning. Mange av språk- og litteraturforskerne ved fakultetet har gjennomgått tekstlabkurs, og vi har inntrykk av at tilbudet brukes av mange. Siden forrige bulletin har det skjedd noe gledelige: Fakultetet har bestemt at Tekstlaben skal tilgodesees med en fast førsteamanuensisstilling. Dermed skulle ansatte og studenter være sikret et godt tilbud også i fremtiden.



Tekstlabens tilbud

Tekstlaboratoriet disponerer en lang rekke tekster som er elektronisk tilgjengelige. Tilbudet på norsk og engelsk er stort, her har vi en god del romaner, lovtekster og avistekster, i tillegg til noen andre typer tekster og noen ferdigsamlede korpora. Vi har også en del elektroniske tekster og ordbøker for andre språk, som albansk, bulgarsk, dansk, estisk, finsk, fransk, gresk, gælisk, italiensk, japansk, kinesisk, latin, litauisk, malayisk, nederlandsk, portugisisk, russisk, serbo-kroatisk, spansk, svensk, tibetansk, tsjekkisk, tyrkisk, tysk og walisisk. Tekstene er av forskjellig type, noen er skjønnlitterære, andre sakprosa, og noen inneholder grammatisk merking, mens andre ikke gjør det. Noen av

tekstene er flerspråklige (parallellkorpus), slik at de inneholder originale og oversatte tekster mellom språk. Utvalget utvides stadig. Særlig ser Tekstlaben det som viktig å øke mengden av norske tekster, men vi prøver også å innhente tekster fra andre språk i den utstrekning vi har grunn til å tro at det er behov for dem. Dessuten kan vi være behjelpelige med å få tak i spesifikke tekster, hjelpe til med å scanne dem inn osv.

I tillegg til tekstene har Tekstlaben en del programvare som er nyttig for å søke i tekstene, og en del programvare som kan være interessant og nyttig for språk- og litteraturforskere mer generelt. Dette gjelder f.eks. WordNet, som inneholder et korpus av engelske leksemer relatert semantisk til hverandre som synonymer, antonymer, hyponymer, antonymer, polysemer osv. Vi har også avanserte søkeprogrammer for PC og Macintosh, som Tact, Word Cruncher, Conc, Pandora o.a. Tekstlaben holder jevnlig kurs i disse programmene, men kan også demonstrere dem individuelt når det er interesse for det.

Tekstlaboratoriets tilbud er tilgjengelig over nettet for brukerne, dvs. vitenskapelig ansatte og høyeregradsstudenter. For at man skal kunne benytte seg av det, må man ha en konto på Hedda. Det har man rett til hvis man er ansatt eller hovedfagsstudent. I tillegg må man fylle ut et skjema hvor man lover å bruke materialet til ikke-kommersielle formål.

Tekstene er lagret på en disk knyttet til fakultetets filtjener Hedda. Se etter filer med "Lesmeg" på disken. Der finnes det nyttig og kortfattet informasjon om de andre filene.

- Slik nås disken fra Macintosh: Hent "Velger" fra eplemenyen. Velg sone HF og symbolet for AppleShare. Velg så MacHedda, og klikk OK. Skriv inn brukernavn og passord, og klikk OK. Velg MacTekstlab, og klikk OK. Nå kommer tekstlabdisken (MacTekstlab) opp på skrivebordet ditt, og da er det bare å åpne den.

- Slik nås disken fra PC: Med et passende tekstbehandlingsprogram legger du følgende linje til fila C:\nfs\drives.bat:

NET USE G: hedda:/hf/hedda/tekstlab/pc
Lagre og start maskinen på nytt. Nå er G: definert som tekstlabdisk.

- Slik definerer du ECI-tekstene for PC: Med et passende tekstbehandlingsprogram legger du følgende linje til fila C:\nfs\drives.bat:

NET USE I: hedda:/hf/hedda/cd/eci-mc1/mnt /cd (NB!
Mellomrom etter mnt)

Lagre og start maskinen på nytt. Nå er I: definert som ECI-område.

Spørsmål om oppkobling mot tekstlabdisken kan rettes til Tekstlaben eller IT-ansvarlige ved instituttene.

Tekstlaboratoriet, HF, UiO, P.b. 1102 Blindern, 0317 Oslo

Faks: 22 85 69 19

Janne Bondi Johannessen (forsker) Helge Hauglin (ingeniør)

Rom B 152, P.A. Munchs hus

Rom B 150, P.A. Munchs hus

Telefon: 22 85 68 14

Telefon: 22 85 78 35

E-post: jannebj@hedda.uio.no

E-post: helgeha@hedda.uio.no

tekstlab@hedda.uio.no

URL: <http://hf.uio.no/tekstlab/>

Ansvarlig for Nyhetsbulletinen: Janne Bondi Johannessen