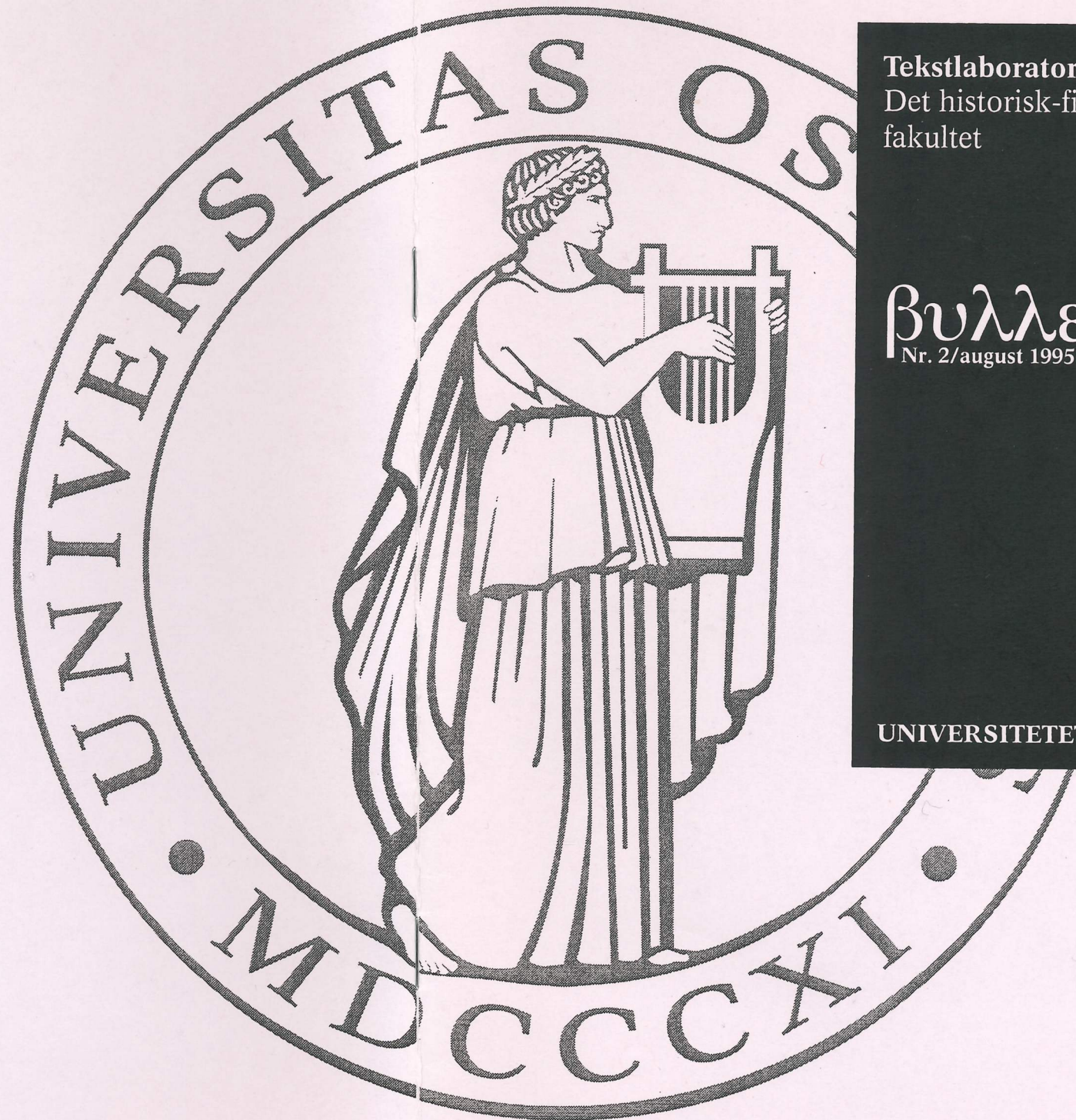


Tekstlaboratoriet

Det historisk-filosofiske fakultet
Boks 1102 Blindern
0317 Oslo

Innhold

- Gjest i dette nummeret
- Oxford English Dictionary i elektronisk form
- Konferanse i Oslo
- Seminar i Princeton
- Tekstlaben for hovedfagsstudenter
- Kort om Tekstlaboratoriet

**Tekstlaboratoriet**

Det historisk-filosofiske
fakultet

ΒΥΛΛΕΤΙΝ
Nr. 2/august 1995 • 3. årgang

UNIVERSITETET I OSLO

Nyhets-

βυλλετιν

fra Tekstlaboratoriet, HF, UIO

3. årgang, nummer 2, august 1995

Innhold

- Gjest i dette nummer
 - Oxford English Dictionary i elektronisk form
 - Konferanse i Oslo
 - Seminar i Princeton
 - Tekstlaben for hovedfagsstudenter
 - Kort om Tekstlaboratoriet (Tekstlabens fremtid og Tekstlabens tilbud)
-

Gjest i dette nummer

Navn: Kjetil Strand
Stilling: Universitetsstipendiat
ved Institutt for lingvistiske fag
Sivilstatus: Gift, to døtre på
4 og 18 år
Hobbyer: Håndball og sjakk



I Tekstlabens nyhetsbulletin starter vi i dag en ny serie med gjester. Hvert nummer vil presentere en gjest (en forsker ved HF-fakultetet) som har hatt hatt glede av Tekstlabens tjenester. Kjetil Strand er en av de etterhvert mange som synes Tekstlabens tilbud har beriket forskningsarbeidet, og han er Nyhetsbulletinens første gjest.

-Kan du kort fortelle hva doktoravhandlingen din dreier seg om?

-Jeg er interessert i substantivfraser i bestemt form (f.eks. *den store mannen* eller *ølprisene*), og hva slags "anker" de har i teksten. Det er nemlig slik at fraser av denne typen ikke kan brukes med mindre de allerede er forankret et sted, enten i teksten eller mer generelt i den ikke-språklige konteksten. For å illustrere med et eksempel: Ordet *mannen* i setningen *Mannen gikk nedover gata* er i bestemt form, og må være forankret et eller annet sted. Han kan for eksempel være omtalt tidligere i teksten gjennom frasen *en rar type* i *Fru Kristensen så en rar type ved butikken i går*.

-Hva slags nytte har du hatt av Tekstlaben?

-I arbeidet med doktoravhandlingen er det viktig for meg å kunne studere større mengder med autentisk tekst. Tekstlaben har jo mye elektronisk tilgjengelig tekst. Jeg har valgt å bruke avistekster, fordi de er relativt varierte. Derfor har jeg tatt for meg et korpus av Aftenposten-tekster, nærmere bestemt hele aftennummeret 10/1-94, som inneholder 69 artikler. Jeg har funnet fram til alle forekomstene av bestemte substantivfraser, altså de som inneholder substantiver i bestemt form - som slutter på -en, -et, -a, eller -ene. Det ble til sammen 1805 stykker (eller 7,7 % av den samlede teksten).

-Har du funnet noe av interesse?

-Åja, det er flere ting som er interessante ved denne undersøkelsen. Én ting er at jeg har funnet at de absolutt aller fleste av disse frasene er forankret i teksten selv. Det er altså forsvinnende få, faktisk mindre enn to prosent, eksempler av typen *seksdagerskrigen* eller *ozonlaget*, som ikke trenger noen nærmere introduksjon. Når man vet at mange moderne hermeneutikere og retorikere hevder at språket står i veien for forståelsen av teksten, er dette viktig kunnskap. Det er altså snarere slik at mye eksplisitt informasjon bibringes av språket.

Men det som egentlig opptar meg mest, er beregnbarheten. Definitte fraser beskriver jo unikt eksisterende objekter. Kan man si noe systematisk om på hvilken måte definitive fraser får sin unikheter? Hvordan hentes kilden til unikheten?

Til å hjelpe meg har jeg laget en enkel database med Filemaker på Mac'en min. Jeg har laget en post for hver eneste definitive frase (og noen ganger flere, hvis frasen har mer enn ett anker). Jeg har undersøkt en lang rekke punkter for hver frase. For det første har jeg delt inn forholdet mellom frase og anker i forskjellige klasser av relasjoner, f.eks. identitet (som i ankeret *en mann* og frasen *mannen*), synonymi (som i *en maskin...apparatet*), hyponymi (*en rose...blomsten*), parafrase (*en ungarer...den gamle mannen*), medlem-mengde (*bruden...paret*),

possesivitet (*Per..hodet sitt*), event-argument (*mannen ...som jeg så klokka fem*), idiom uten forankring (som i *løpet av*) osv. Av disse relasjonene står identitetsrelasjonen for 22% av tilfellene, noe som var å vente. Mer interessant er det at parafraserelasjonen, hvor én leksikalsk frase er forankret i en helt annen leksikalsk frase, er brukt i nærmere 18% av tilfellene.

Videre har jeg sett på hvor i teksten forankringen har vært. Ofte (20%) er forankringen inkludert i en frase og modifierer den definite frasen (som i *grekerne...på Rhodos*). Ellers er selvfølgelig anafori (dvs. at forankringen forekommer et eller annet sted tidligere i teksten) svært vanlig, slik forankring skjer i 65% av tilfellene, mens katafori (hvor ankeret forekommer senere) er heller sjeldent (2,5%). Det som er mest påfallende, er som nevnt at en forankring utenfor teksten er svært uvanlig (mindre enn 2%).

-Er virkelig forholdet mellom bestemt frase og anker beregnbart?

- Dessverre må jeg si at resultatene så langt er skuffende når det gjelder beregnbarhet. Avstandene mellom definitiv frase og anker varierer fryktelig. Men de er likevel mindre enn andre forskere har anslått. Gjennomsnittlig er det 2,8 setninger mellom dem, og under 35 ord.

En overraskende oppdagelse er det at svært få av ankerne er i fokus (dvs. svært nærværende i talernes og lytternes oppmerksomhet), i sterk kontrast til det som er tilfelle med ankerne til pronomener, som jeg også har arbeidet med. Dette er kanskje ikke så rart, siden grunnen til at man velger en leksikalsk snarere enn en pronominal uttryksform nettopp kan være at man behøver å gjenta eksplisitt hvem eller hva man refererer til.

Men selv om det kanskje er vanskelig å generalisere alle de definitive frasene på samme måte, har jeg et håp om at det kan finnes generaliseringer for undergrupper av tilfeller. For eksempel synes jeg at jeg ser en sterk tendens til at de forskjellige relasjonsklassene knytter

seg automatisk til ulik grad av nylighet og fokus i teksten. Det ser ut til at possessive relasjoner har veldig nylige eller høyfokuserende ankere. Det er vanlig at relasjonen finnes innenfor samme setning, for eksempel at objektet er forankret i subjektet: ... så *jeg holdt på å brette nesen (min)*. Videre synes parafrase-klassen å ha et sterkt fokusert anker, som når *disse helt spesielle ishockeyspillerne* viser tilbake til *de*, som igjen viser tilbake til *mange VIF-spillere*.

Jeg håper også at det kan gå an å finne en systematikk som bygger på ulike leksikalsk-semantiske klasser. Det er eksempelvis tydelig at klassen av "vær-leksemer", som inneholder ord som *sola*, *grålysningen*, *temperaturen*, *varmen*, *kulda*, og *snøen*, trenger en forankring av argument-event-typen. Her er det altså tid og sted som må være kjent for at man skal kunne benytte og forstå disse definitive frasene. Andre typer leksemer som trenger forankring av denne typen er *kriminalitetet* og *ølprisen*. Jeg har også funnet at klasser som humanklassen og institusjonsklassen (med bl.a. butikker og skoler) ofte deltar i possessivrelasjonen.

Det er forøvrig også et nytteaspekt ved denne forskningen. Hvis jeg finner den systematikken jeg håper på, kan man bruke resultatet til å beregne ulike ting. F.eks. kan man automatisk generere den semantiske representasjonen av (avis)tekster, noe som kan være svært nyttig ved informasjonsgjenfinning.

-Hva er planene videre?

- Jeg er faktisk på vei til Amerika nå. Jeg skal være et år ved Stanford University, og håper å få forsket mye. Det er jo et svært inspirerende miljø, der som her, får jeg si, med hundrevis av arbeidsgrupper og seminarer. Forhåpentlig får jeg presentert mine ting og utvekslet erfaringer og ideer.

Oxford English Dictionary i elektronisk form



Tekstlaboratoriet er i ferd med å kjøpe inn den elektronisk utgaven av Oxford English Dictionary (OED). Denne ordboken er av mange ansett som det ypperste innen ordbøker. Den er svært omfangsrik, og inneholder et vell av informasjon om oppslagsordene, ikke minst når de er påtruffet første gang i litteraturen. For mange er OED ekspertene som får det siste ordet når en etymologi eller definisjon diskuteres. Likevel skal man være forsiktig med å bruke den helt ukritisk, for selv OEDs redaktører hadde sine personlige preferanser. I 1888-utgaven skrev redaktør Murray at de viktigste kildene var "alle de store engelske forfattere fra alle tidsepoker". Likevel har tydeligvis noen forfattere, særlig Shakespeare, vært viktigere enn andre, slik at retningslinjene om at alle ord funnet i litteraturen skulle tas med, bare gjaldt utvalgte forfattere. For eksempel er nærmere 50 ord fra forfatteren Nashe ikke tatt med. Disse inkluderer ord som *bestialness*, *bastardship*, *absurdify*, *misinformedly*, og andre som typisk ville vært med ellers.

Når den elektroniske utgave er vel i hus, er alle invitert til å komme på en demonstrasjon. Da har også Stig Johansson fra IBA sagt seg villig til å holde en innledning hvor OED ikke bare roses, men kanskje også rises. Den elektroniske utgave er imponerende, og man har mulighet til å slå opp kryssreferanser på svært mange måter. Den blir tilgjengelig over nettet.

Konferanse i Oslo



I april inviterte Institutt for britiske og amerikanske studier til The Second Nordic Symposium on Text-based Contrastive Studies. Det var god deltagelse fra både Norden og resten av Europa. Etter å ha overvært to dager med forelesninger, var nok alle deltagerne overbevist

om at parallelle korpora, altså korpora med tekster fra flere språk, gjerne oversatt til hverandre, kan gi språkforskningen en ny dimensjon. Vi kan ikke nevne alle foredragene her, man vil konsentrere oss om noen få. Vi håper disse smakebitene kan gi litt inspirasjon til andre som kune tenke seg å begynne med elektroniske korpusstudier.

Josef Schmied fra TU Chemnitz-Zwickau, Tyskland, hadde samarbeidet med **Hildegard Schäffler** om en korpusbasert kontrastiv analyse og oversettelsesstudier, og brukt engelsk-tyske oversettelser. De hadde studert det etter hvert velkjente fenomenet "translationese" (skal vi oversette det til oversettersk?), og funnet to interessante undergrupper. Den kjente typen oversettersk er jo den hvor kildepråket slår igjennom til målpråket. Denne typen trenger man ikke parallellkorpora for å oppdage. Schmied illustrerte med en oversettelse av en turistbrosjyre:

- (1a) Many local authorities and conservation bodies have experienced staff available whose task it is to help visitors make the most of Scotland's outdoor environment.
- (1b) Viele örtliche Behörden und Landschaftsschutzorganisationen verfügen über erfahrene Mitarbeiter, deren aufgabe darin besteht, besuchern dabei zu helfen, das meiste aus Schottlands freier Natur zu machen.

I eksemplet slår engelsken igjennom ved at the most er oversatt til das meiste, istedenfor das beste, som det jo heter i det tyske uttrykket.

Den andre typen oversettersk som Schmied redegjorde for, er mer interessant fra et korpusperspektiv, fordi det nettopp er kvantitet man må undersøke. Denne typen kan karakteriseres ved at det er målpråket som slår igjennom ved oversettelse til målpråket, eller med andre ord: Oversetterens kunnskap om målpråket gjør at oversettelsen blir mer fjern fra kildepråket enn nødvendig ut fra målpråkets normer.

F.eks. ser det ut til at det er en språk-kulturell forskjell mellom tysk og

engelsk, hvor tysk er mer direkte, mens det i engelsk er vanlig å pakke inn en proposisjon i ymse modale verb, konstuksjoner og partikler. Dette illustreres i de følgende eksemplene:

- (2a) ... and you may prefer to take advantage of a scauffeurpunt.
 (2b) Sie können auch Stosskähne mit einem Steuermann mieten.
 (3a) ... and whilst we do not all go round singing, acting or composing poetry, it does perhaps hint at the importance that culture has for us.
 (3b) ... und wenn wir auch nicht alle singend, schauspielernd oder dichtend umhergehen, wird einem so doch der Stellenwert der Kultur in Wales bewusst.

Hvis det er riktig at det altså finnes en type oversettersk hvor målspåket påvirker målspåket, skulle man vente å finne mer av den direkte typen språk i tyske oversettelser enn i tyske originaltekster. Og dette viser seg faktisk i noen grad å stemme ut fra det materialet Schied og Schäffler har. I en undersøkelse fant de at mens de originale tyske tekstene hadde 103 tilfeller av indirekte språkbruk, hadd de oversatte færre, nemlig 98 (de engelske hadde 134). Men den store forskjellen var at mens den indirekte språkbruken i de oversatte tyske tekstene ble realisert gjennom 38 tilfeller av adverbialer og partikler (den engelske originalen hadde 33), hadde de originale tyske tekstene hele 78 slike. Det er opplagt at det bare er gjennom korpusstudier at man kan finne fram til denne typen data.

Tove Jacobsen fra Universitetet i Bergen snakket om forskjellen på bevegelsesverb i norsk og engelsk på den ene siden og fransk på den andre. Denne forskjellen kunne beskrives skjematisk:

Norsk		Fransk	
Måte, Bevegelse + Retning		Retning + Måte, Bevegelse	
Verb	Partikkel, Preposisjon	Verb	Prep.frase, Adv.

Eksempler:

fare av sted:	partir en trombe	(Toget fór forbi - Le train est passé <u>à toute vitesse</u>)
galoppere:	galoper	(Han kom galopperende - Il est arrivé <u>au galop</u>)
sjokke:		(Hun sjokket ut av døren - Elle est sortie <u>d'un pas traînant</u>)
skynde:		(skynde seg opp trappen - monter <u>vivement</u> l'escalier)
svømme:		(svømme over elven - traverse la rivière <u>à la nage</u>)
tusle:		(Han tuslet ut - Il est sorti <u>à pas feutrés</u>)

Forskjellen mellom språkene kommer svært godt fram ved studier av parallellkorpora. Man blir oppmerksom på at tekster i de to språkene vektlegger forskjellige aspekter, nettopp fordi de grammatiske konstruksjonene er forskjellige. Således vil de franske tekstene ofte unnlate å uttrykke eksplisitt det som må realiseres som preposisjonsfrase eller adverb, nemlig måte og bevegelse, så sant man kan utlede det av sammenhengen. I eksemplet nedenfor kan man utlede at rytterne kom ridende, derfor er ikke måten de kom på, tatt med:

The horsemen rode into the yard
 Les cavaliers sont entrés dans la cour.

Den typen kunnskap man får av å studere parallellkorpora på denne måten er svært nyttig når man skal lage tospråklige ordbøker. I tillegg til mulige oversettelser, kan man ved å studere korpora finne hvilke som er vanlige. Slik sett burde kanskje ordbøkene skille mellom de beskrivende oversettelsene, som er mulige, men som ofte kan utelates

(som ved oversettelsen av *fare* over), og de som inneholder faste uttrykk (som oversettelsen av *svømme*).

Diana Santos, INESC, Lisboa, og IBA, Universitetet i Oslo, hadde studert tempus- og aspektsystemene i engelsk og portugisisk på bakgrunn av parallelle korpora. Slik kom hun fram til en semantisk typologi for de forskjellige tempus- og aspekt-kategoriene og oversettelse av dem.

Hun viste at engelsk *Simple Past* oftest blir oversatt med *Imperfeito* eller *Perfeito*, og bare sjelden med *Infinitivo*, *Gerúndio*, *Presente* eller andre. Tilsvarende blir *Imperfeito* oftest oversatt til *Simple Past*, og det samme gjør *Perfeito*. Likevel bli oversettelsene bare omtrentlige, fordi en god del informasjon som følger med de portugisiske tempuskategoriene ikke impliseres i de engelske. For eksempel gir *Imperfeito* gir et internt perspektiv (samtidighet) på handlingen, dette blir rett og slett borte i den engelske oversettelsen:

Portugisisk: Ia de olhos baixos, temendo encontrar alguém conhecido, alguém com quem fosse obrigado a trocar palavras.

Engelsk: He walked with his eyes down, afraid of meeting anyone he knew, anyone with whom he would be obliged to exchange a few words.

På tilsvarende måte vil oversettelse fra engelsk til portugisisk uvegerlig måtte legge til informasjon som ikke impliseres i originalen.



Seminar i Princeton

I juni deltok Tekstlaboratoriets to ansatte ved et to ukers seminar ved Princeton University, New Jersey, USA. Det var Center for Electronic

Texts in the Humanities som holdt sitt årlige sommerseminar, og vi deltok sammen med 58 andre fra forskjellige deler av verden. Seminaret var arrangert med både plenumsesjoner og gruppesesjoner for spesielt interesserte. Tekstlabens forsker deltok på gruppen for Textual Analysis, mens ingeniøren var på gruppen for SGML/TEI-koding av tekster.

SGML/TEI-koding er svært i vinden for tiden, det dreier seg om systemer man kan bruke for å kode forskjellig informasjon på en enhetlig måte. Vi tenker oss at det skal brukes for å merke Tekstlabens elektroniske tekster med bibliografisk og annen nyttig informasjon. Som det jo er nå, er det svært tungvint å finne ut hvilken tekst man er inne i når man først har åpnet en tekstfil. Tekstanalyse er selvfølgelig det som humanister flest lettest kan forholde seg til. Hvordan man kan bruke ulike tekstsøkeprogrammer for å finne ulike ting er relevant for selve kjernen i mye språk- og litteraturforskning. Lærerne Susan Hockey (Princeton) og Willard McCarthy (Toronto) hadde mange nyttige tips.

Plenumsesjonene ble holdt av inviterte forelesere fra Europa og Amerika. Det ble holdt en plenumsesjon for hvert emne som det fantes gruppesesjoner i, slik at alle fikk et innblikk i de forskjellige emnene. Det var forelesninger i verktøy for historisk analyse, hvordan starte og vedlikeholde et elektronisk tekstsenter, litterær redigering, og hypertekst i humaniora, i tillegg til tekstanalyse og SGML/TEI-koding. Videre var det gjesteforelesere som demonstrerte interessant programvare av ulike slag til bruk i humaniora.

Det mest imponerende var nok Piero-prosjektet som var utviklet ved Princeton, som et samarbeid mellom instituttene for kunsthistorie og informatikk. Et helt kapell med fresker fra gulv til tak av Piero della Francesca var blitt fotografert og digitalisert. Man kunne "gå" gjennom kapellet i alle krinkler og kroker, og opp og ned, og stille seg opp i luften noen meter over bakken for å kikke nærmere på bildene høyt oppe på veggen. Dessuten kunne man holde en lyskaster i hånden, og sveipe over det man måtte være interessert i. Lyset fra vinduene forandret seg etter hvordan solen sto på himmelen, og dette fikk

selvsagt konsekvenser for lyset på bildene. Det er fantastisk hvilke perspektiver denne typen prosjekter gir. I tilfellet med Piero-bildene kan folk nå for første gang få se på nært hold bilder som bare kunstneren selv og restaureringsarbeiderne tidligere har fått se.

Perseus-prosjektet som ble presentert, var også svært imponerende. Her har man samlet store deler av klassiske greske tekster, og bundet dem sammen med bilder og leksikon, slik at hvis man klikker på ordene i den originale teksten, får man oversettelse og morfologisk informasjon frem. Klikker man et annet sted, får man bilder fra stedene og tingene som er beskrevet i teksten.

I tillegg til alle forelesningene og gruppene var det også meningen at seminardeltagerne skulle arbeide med individuelle prosjekter. Noen av disse ble så presentert de siste dagene for resten av deltagerne og noen innbudte gjester. Tekstlabens ingeniør kodet et par norske romaner på SGML-vis, mens forskerens prosjekt var å undersøke distribusjonen til norske personlige pronomener ved hjelp av Tact-programmet og en norsk roman. Pronomen-prosjektet ble presentert, med en viss suksess, som antagelig for en stor del skyldtes at publikum syntes det var morsomt å høre et skandinavisk språk til en avveksling.

Alt i alt var seminaret veldig nyttig. Vi lærte mye nytt, og fikk et godt inntrykk av hva som rører seg i skjæringspunktet mellom teknologi og humaniora i den store verden. Det er tydelig at Tekstlaboratoriet er et ektefødt barn av vår tid, og har mange brødre og søstre rundt i verden, og med enda mange flere i vente.

Tekstlaben for hovedfagsstudenter



I august ble Tekstlaben for første gang presentert for hovedfagsstudenter, da vi fikk holde en orientering med øvelser på Hovedfagsstartkurset. Dette var vi glad for å kunne gjøre, for det er

helt sikkert at mange hovedfagsstudenter kan dra nytte av vårt tilbud av elektroniske tekster, programvare, og muligheter til å hjelpe folk til å bestille eller scanne inn elektroniske tekster. Vi fikk litt for kort tid til å få gjennomgått alt vi ønsket - forhåpentlig får vi litt mer tid neste gang. De aller fleste studentene var heldigvis entusiastiske.

Kort om Tekstlaboratoriet



Tekstlabens fremtid

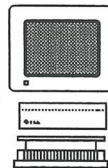
Tekstlaben har i dag to ansatte, en forsker (dr. philos.) og ingeniør (cand. scient.). Ingeniøren er ansatt av fakultetet og stillingen er permanent. Forskerstillingen er knyttet til prosjektet Elektroniske verktøy i språkforskningen, som er finansiert av Norges forskningsråd og som utløper om et års tid. På tross av tittelen er ikke dette en ren forskerstilling. Halve arbeidstiden blir brukt til kursvirksomhet ved og drift av Tekstlaben. Tekstlaben har bestått i sin nåværende form i to år, og både vi og Tekstlabens styringsgruppe er godt fornøyd med resultatet så langt. Mange av språk- og litteraturforskerne ved fakultetet har gjennomgått tekstlabkurs, og vi har inntrykk av at tilbudet brukes av mange.

Vi mener at driften i denne perioden har vist behovet for og nytten av Tekstlaben. Nå er det viktig at Tekstlaben fortsetter å være et kraftsenter som tilbyr elektroniske tekster og programvare av ulike slag til gamle og nye brukere, og at vi passer på å fornye og utvide vårt utvalg av hjelpemidler som kan være til nytte for fakultetets forskere, det være seg tekster eller programmer. I dette inngår også at vi selv utvikler og driver forskning på eller ved hjelp av Tekstlabens verktøy.

For at denne virksomheten også skal fortsette etter at NFR-prosjektet tar slutt neste år, må virksomheten få et mer permanent preg og Tekstlaben tilføres en fast førsteamanuensisstilling ved siden av ingeniørstillingen. Beslutningen om dette ligger hos fakultetet. At brukerne synliggjør sin entusiasme i denne sammenheng, vil ikke skade!



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 hjelpelige 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

Ansvarlig for Nyhetsbulletinen: Janne Bondi Johannessen