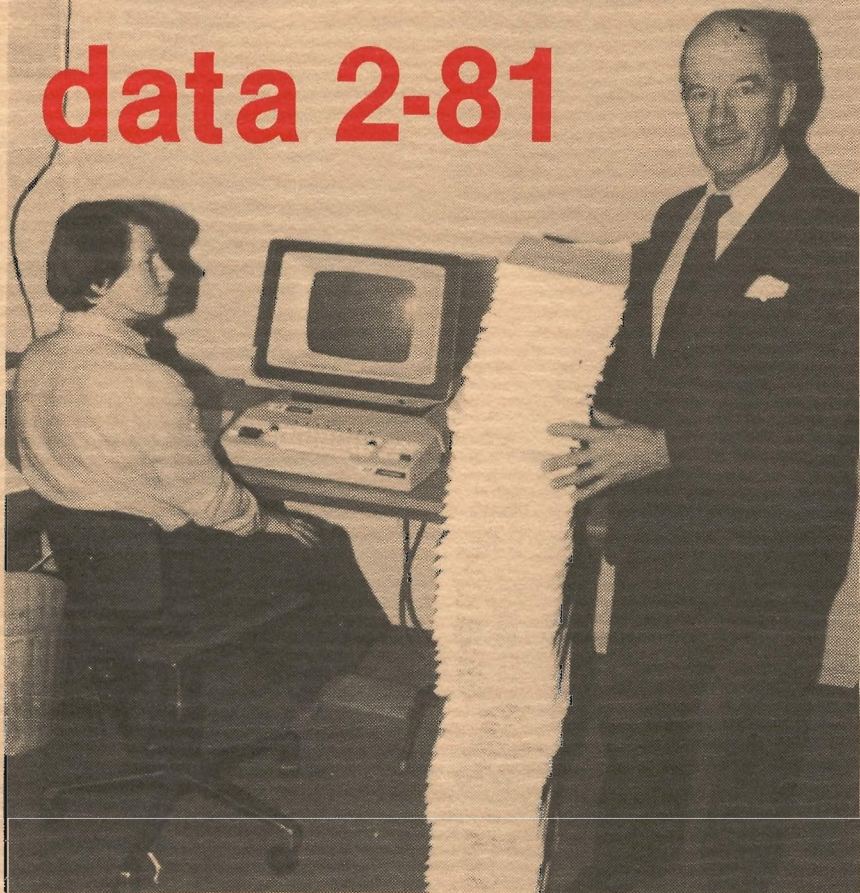


# humanistiske data 2-81



ARTIKLER  
RAPPORTER  
MELDINGER  
SUMMARY

NAVFs EDB-senter  
for humanistisk  
forskning

---

The Norwegian  
Computing Centre for  
the Humanities.

## SENTERETS RAPPORTSERIE

RAPPORT nr. 1. *EDB i gjenstandsfagene*. Rapport fra en konferanse i Bergen, 18. og 19. april 1978. September 1978. Pris kr. 15,

RAPPORT nr. 2. *Et norsk datamaskinelt tekstkorpus*. Rapport fra en konferanse i Bergen, 19. og 20. oktober 1978. Febr. 1979. Pris kr. 20,

RAPPORT nr. 3. *Rapport fra den nasjonale konferanse om EDB i språk og litteraturforskning*, 4. og 5. januar 1979. Mars 1979. Pris kr. 25,

RAPPORT nr. 4. *Oppbygging av EDB-katalog for folkemusea i Hordaland og Kulturgeografisk registrering på Vestlandet*. April 1978. 2. opptrykk oktober 1979. ISBN 82-7283-000-0. Pris kr. 15,-.

RAPPORT nr. 5. *Rapport fra NKKM's EDB-komite*. August 1979. ISBN 82-7283-001-9. Pris kr. 15,-.

RAPPORT nr. 6. *Prøveprosjekt med EDB ved Norsk Folkemuseum*. Oktober 1979. ISBN 82-7283-002-7. Pris kr. 15,-.

RAPPORT nr. 7. *Ivar Fønnes: Norsk landbruksordbok. Prosjektrapport om databehandling og tilrettelegging for trykking*. September 1979. ISBN 82-7283-008-6. Pris kr. 25,-.

RAPPORT nr. 8. *SEFRAK. Rapport frå prøveprosjekt for databehandling av kulturminneregisteret*. Oktober 1979. ISBN 82-7283-003-5. Pris kr. 20,

RAPPORT nr. 9. *Jostein H. Hauge og Sigbjørn Århus: Dataregistrering i humanistiske fag med vekt på optisk lesing*. August 1978. 2. opptrykk oktober 1979. ISBN 82-7283-004-3. Pris kr. 15,-.

RAPPORT nr. 10. *Roald Skarsten: Innføring i SPSS for humanister*. November 1977. 2. opptrykk november 1979. ISBN 82-7283-005-1. Pris kr. 15,-.

RAPPORT nr. 11. *Jostein H. Hauge og Knut Hofland: Rapport fra 4 konferanser i USA sommeren 1979*. The 17th Annual Meeting of Computational Linguistics. La Jolla Conference on Cognitive Science. The fourth International Conference on Computers in the Humanities. Data Bases in the Humanities and Social Sciences. November 1979. ISBN 82-7283-007-8.

RAPPORT nr. 12. *EDB og manuskriptregistraturer*. Oktober 1977. 2. opptrykk november 1979. ISBN 82-7283-009-4. Pris kr. 20,-.

*fortsettelse 3. omslagsside.*

# humanistiske data 2-81

---

NAVFs EDB-senter for  
humanistisk forskning

---

The Norwegian Computing  
Centre for the humanities

---



NORGES  
ALMENVITENSKAPELIGE  
FORSKNINGSRÅD

*Humanistiske Data* blir utgitt av NAVFs EDB-senter for humanistisk forskning. Redaksjon: Jostein H. Hauge, Svein Skotheim. Bladet kan bestilles ved henvendelse til Senterets adresse: Harald Hårfagesgt. 31, Boks 53, 5014 Bergen-Universitetet. Tlf. (05) 210040, linje 2956.

Artikler, rapporter, meldinger mottas. Redaksjonen avsluttet 1.september.

---

*Humanistiske Data* is published by The Norwegian Computing Centre for the Humanities. Editors: Jostein H. Hauge, Svein Skotheim. The journal can be ordered from the adress mentioned above. Contributions are welcome.



*Informasjonsarbeid utgjør en viktig del av Senterets arbeid. Det gjelder publisering, kurs- og konferansevirksomhet, konsulenttjenester m.v. Som et ledd i opptrapping av den utadrettede informasjonsvirksomheten er Svein Skotheim (bildet) tilsatt som informasjonskonsulent i halv stilling. Redigering av *Humanistiske Data* vil bli en av hans mest sentrale oppgaver, og bladet vil fra 1982 utkomme med 4 nummer pr. år med et opplag rundt 1500.*

Medarbeidere i dette nummer:

*Knut Fintoft*, professor, Lingvistisk institutt, Universitetet i Trondheim

*Ivar Fønnes*, amanuensis, HF-data, Universitetet i Oslo.

*Knut Hofland*, førstekonsulent, NAVFs EDB-senter for humanistisk forskning.

*Eirik Lien*, konsulent, EDB-tjenesten for filologiske fag, Universitetet i Trondheim

*Øystein Reigem*, konsulent, NAVFs EDB-senter for humanistisk forskning

*Åse Støgdard*, univ.bibliotekar, UB, Bergen.

*Per Vestbøstad*, konsulent, Norsk tekstarkiv

*Sigbjørn Århus*, førstekonsulent, NAVFs EDB-senter for humanistisk forskning

---

## **FORSIDETEKST**

*Slik kan hundreårsutgaven av Ibsens samlede verker se ut når den er lagt inn på EDB. Professor Harald Noreng avslutter i høst EDB-siden av det store Ibsen-prosjektet. *Humanistiske Data* vil komme tilbake til dette prosjektet i senere numre.*

*Fotosats i kommunikasjon med Univac 1100/82*

*Sats og trykk: Bergmanns Boktrykkeri A/S, Bergen*

## **Innhold**

<i>Artikler</i>	side
Fonetikk og statistikk. <i>Knut Fintoft</i> .....	4
EDB og humaniora. <i>Jostein H. Hauge</i> .....	10
Intervju med <i>Ivar Fønnes</i> .....	12
 <i>Rapporter</i>	
National Computer Conference .....	15
The Fifth International Conference on Computers in the Humanities .....	16
Joint Conference on Easier and more Productive Use of Computing Systems .....	20
On the Possibilities and Limit of the Computer in producing and publishing Dictionaries .....	22
Nordisk forskerkurs om multivariate metoder i arkeologi .....	28
Nordisk forskerkurs i språklig databehandling .....	29
NordDATA .....	31
The 6th. International Congress of Applied Linguistics .....	33
Tre glade hospitanter .....	34
Utviklingsseminar .....	35
 <i>Meldinger</i>	
Forskere knyttet til Senteret .....	37
7 konferanser i 1981/82 .....	37
Kurs for museumsansatte .....	41
EDB-tjenestene ved universitetene .....	41
BIBNETT .....	43
Nye rapporter fra Senteret .....	44
Summary .....	46

# Fonetikk og statistikk

*Knut Fintoft*

En del vil kanskje i første omgang tro at statistikk og databehandling må være meget fjernt fra og uvesentlig i et fag som fonetikk. Det kommer vel av at mange har en feilaktig oppfatning av faget. De fleste har kommet i kontakt med faget i skolen, i forbindelse med uttale av fremmedspråk. Dette er en del av faget, en del av den anvendte fonetikk, men fonetikk er mye mer.

Jeg vil i det følgende vise ved eksempler noen typer av problemer en arbeider med innenfor faget. Jeg vil presentere en del spørsmål og la mange stå ubesvart. Hensikten med denne presentasjonen er bare å påpeke nødvendigheten av statistikk og EDB i faget fonetikk.

Fonetikk grenser mot mange fag og disse fags metoder vil derfor prege de ulike sider av faget fonetikk. Når det er snakk om innsamling av data, metode, databehandling og statistikk må en spesielt nevne tilknytningen til følgende fagområder, supplert med noen stikkord:

*teknisk-naturvitenskapelige fag* — måling av fysiske størrelser (måleusikkerhet, måleobjektets ubestemthet)

*psykologi* — persepsjonstester (validitet, reliabilitet)

*sosiologi* — fonetiske forhold betinget av sosiale faktorer

*pedagogikk/medisin* — utprøving av midler

Innsamlingen av data foregår på forskjellige måter innenfor de ulike deler av fonetikken, men uansett området må dagens fonetikere kvantifisere sine observasjoner og dermed benytte en eller annen form for databehandling. Ved Universitetet i Trondheim har vi tatt konsekvensen av det, behandling av data inngår i fagstudiet fonetikk.

La oss se på noen eksempler fra ulike felt av faget.

## **Hypypighet av språklyder og kombinasjoner av språklyder.**

På tilsvarende måte som en er interessert i forekomsten av grafemkombinasjoner og ord, er det naturlig at en også er interessert i tilsvarende ting når det gjelder språklyder. Men det å fremskaffe et slikt lydmateriale er problematisk og arbeidskrevende, og særlig problematisk er det dersom en ønsker en utlistering av de enkelte ord. En må imidlertid ta utgangspunkt i genuin tale — ikke i antakelse om hvorledes trykt tekst blir lest. Erfaring viser at en har atskillig mer assimilasjon og utelatelser i normal tale enn en vanligvis tror. Det første problem en står overfor er identifikasjon og transkripsjon av de enkelte språklyder. Erfaring viser at vi meget ofte uttaler andre lyder enn de vi selv og omgivelsene oppfatter. Det kreves derfor ofte ganske mye av den som transkriberer for å oppdage dette. Vi må også ta stilling til om vi skal foreta en

---

Artikkelen bygger på et foredrag holdt på *Kurs i statistikk for språk- og litteraturforskere*, Oslo 9. — 12. juni 1980.

undersøkelse av foner eller av fonemer. Hvis vi velger å undersøke fonemer, må vi på forhånd fastlegge inventaret av fonemsystemet. Skal f.eks. [ t ] betraktes som ett fonem /t/ eller som en kombinasjon av fonemene /r/ og /t/? Fonemsystemet kan også bli problematisk dersom en ønsker å undersøke både barnespråk og voksespråk. Har en så endelig transkribert materialet, kan det behandles nesten analogt til vanlig ortografi. Det er dog visse problem. I ortografi er ordene skilt fra hverandre. Tilsvarende finnes ikke i tale, lydene følger direkte etter hverandre, vanligvis uten pause mellom ordene. La oss se på noen eksempler, representert ved ortografi og en fonetisk transkripsjon:

far: [ fa:r ]  
fars skjorte [ fa:ʃut:ə ]  
far til Per [ fa:ʃipe:r ]

Ortografisk er det ikke vanskelig å skille ut 'far', men noe tilsvarende er knapt meningsfylt når det gjelder uttale.

### **Bestemmelse av formantfrekvenser.**

Vokalenes formanter — topper i frekvensspekteret — kan med en viss rett betraktes som vokalenes 'fingeravtrykk'. Det har derfor vært av betydning å bestemme de ulike vokalers formantfrekvenser, bl.a. for generering av syntetisk tale. Formantfrekvensene er også av stor betydning i en rekke andre sammenheng som knytter seg til artikulasjon og persepsjon. Det knytter seg en rekke både teoretiske og praktiske problemer til måling av formantfrekvenser. I norsk vil en finne at formantene er avhengig av dialekt, på grunn av ulike vokalkvaliteter og eventuelle ulike fonemsystem. Dessuten er formantfrekvensene avhengig av kjønn, til en viss grad av alder og ikke minst av nabolyd. Det siste fører til at formantfrekvensene til de ulike lyder varierer i løpet av lydens varighet. Det vil si at formantfrekvensene er avhengig av en rekke variabler. Men det kan være formålstjenlig å måle formantfrekvensene i ulike vokaler med ulike kombinasjoner av nabolyd og holde variablene dialekt, kjønn og alder konstant. For de målte verdier vil det være naturlig å angi et mål for den sentrale tendens (f.eks. aritmetisk middelvei) og spredning (f.eks. standardavvik) som deretter eventuelt kan anvendes i videre statistiske analyser.

### **Varighetsforskjell mellom 'lang' og 'kort' vokal.**

Er det f.eks. noen varighetsforskjell mellom /i:/ og /i/? En støter med en gang på et både teoretisk og måleteknisk problem, nemlig definisjon av grensen mellom to språklyder. Normalt er det nemlig en kontinuerlig overgang fra en lyd til en annen, på tilsvarende måte som det er en kontinuerlig overgang fra en bokstav til en annen i vanlig håndskrift. Men selv etter at vi har definert grensene støter vi på en del andre problem. Det viser seg at varigheten av en språklyd er avhengig av nabolydene (egentlig nabolydenes varighet, om de er stemte eller ustemte), av posisjonen i ordet, antall stavelser i ordet, trykkplassering og taletempo. For å besvare spørsmålet vi stilte, vil det være aktuelt å måle vokalvarigheten i minimale par med ulike omgivelser som

f. eks.  
fine pine  
finne pinne

Etter at et visst antall personer har uttalt disse ordene, kan en beregne middelveidene og standardavvikene for de aktuelle vokalvarighetene og deretter benytte standardprosedyrer for å bestemme signifikansen mellom to middelveidier.

### **Midling av EMG-signaler.**

I den fysiologiske delen av fonetikken er en bl.a. interessert i å kartlegge aktiviteten av muskler som er relevante for tale, f.eks. i tunga og i strupehodet. Muskelaktiviteten kan måles ved det elektriske spenningspotensial i muskelen, enten ved hjelp av elektroder som er festet til huden eller ved nålelektroder som stikkes inn i muskelen. Signalet (spenningsvariasjon med tiden) kan ha en viss tendens til å drukne i støy, slik at det er vanskelig å detektere. Støyen er ofte tilfeldig (random noise). Dersom en tenker seg at en adderer flere slike kurver, vil støyen kunne reduseres, mens signalet forsterkes. Jeg skal ikke gå videre inn på dette, men bare nevne at en teknikk som brukes kalles autokorrelasjon. Praktiske løsninger forutsetter at en benytter datamaskin. I en del andre vitenskaper er en også opptatt av signaldeteksjon av svake signaler og i fonetikken benytter en teknikk fra disse vitenskaper.

### **Forskjellen mellom to tidsfunksjoner.**

I en rekke tilfeller vil det være aktuelt å sammenlikne to tidsfunksjoner, f.eks. to tonelagskurver (dvs. grunnfrekvensen som funksjon av tiden). En enkelt registrering fra en person kan f.eks. være som i fig. 1. Kan kurvene for tonelag 1 og 2 sies å være forskjellige? Hvor stor må forskjellen være for at vi skal kunne si at de er forskjellige? Det er kanskje litt drastisk å uttale noe generelt om tonelagskurver ut fra en enkelt registrering. For det første er kurvene avhengig av dialekt, men også innenfor den enkelte dialekt er kurvene noe forskjellige fra gang til gang, både i frekvens og tid, selv om en gjentar de samme ordene. Dessuten er kurvene til en viss grad avhengig av de enkelte lyder. Den del av kurvene som har størst betydning er knyttet til den trykksterke vokalen pluss den etterfølgende konsonant (dvs. /an:/). Kurvene kan normeres i frekvens og tid med den nevnte kurvedel som enhet. De relative posisjoner av maksima og minima (angitt ved middelveid og standardavvik for en rekke kurver) kan nå sammenliknes. Jeg skal ikke gå videre på dette tema, men vil påpeke at den produksjonsmessige forskjell, som vi nå har sett på, ikke nødvendigvis er lik den perseptoriske forskjellen. Dersom vi ønsker å vite hvilken del av kurvene som er perseptorisk mest relevant, må dette undersøkes ved en persepsjonstest.



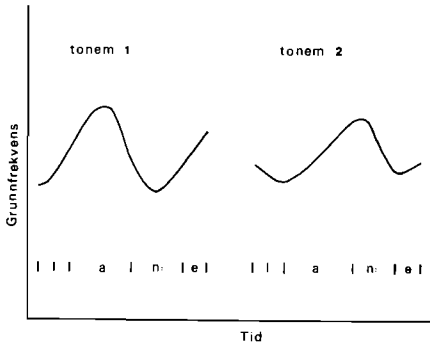


Fig. 1. Eksempel på tonelagskurver for tonem 1 og 2 (landet — lande).

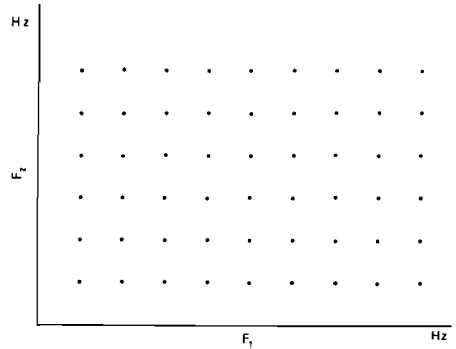


Fig. 2. Kombinasjon av formantfrekvensene F1 og F2 i faste sprang.

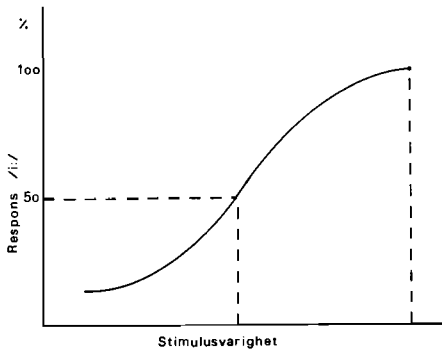


Fig. 3. Antatt prosentvis oppfattelse av /i:/ ved reduksjon av vokalens varighet.

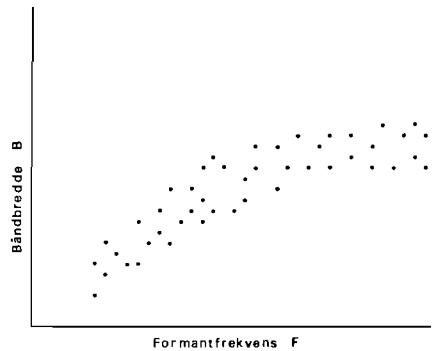


Fig. 4. Sammenhørende verdier for vokalformanters frekvens og båndbredde.

### Persepsjonstest med syntetiske vokaler.

Vi ønsker å vite innenfor hvilke grenser formantfrekvensene må ligge for at vi skal oppfatte en bestemt vokal. Ett sett av formantfrekvenser gir én vokalkvalitet. Vi kan utføre et eksperiment der vi lar første og andre formant (F1 og F2) variere i faste sprang (se fig.2) For hver kombinasjon av F1 og F2 spør vi forsøkspersoner hvilken vokal de oppfatter lyden som. Responsen vil være avhengig av faktorer som antall svarsalternativ (antall fonemer eller antall fonemer som kan være realisert ved hvert stimulus), dialekt (på grunn

av ulik vokalkvalitet) og grunnfrekvensen (dvs. kjønn og alder til den 'person' som har uttalt vokalene). For noen F1-F2 kombinasjoner vil det kanskje bare være én eller to naturlige svarsalternativer, mens det for andre kombinasjoner kan være flere alternativer med ulik vekt. Persepsjonsgrensene kan eksempelvis defineres slik at de avgrenser et område der den aktuelle vokal oppfattes med en viss sannsynlighet.

### **Persepsjon av 'lang' og 'kort' vokal.**

Vil oppfattelsen av f.eks. /i:/ og /i/ være avhengig av vokalens varighet? Er det varigheten av vokalen som er avgjørende for om vi skal oppfatte vokalen som 'lang' eller 'kort'? I tilfelle hvor kort må [ i : ] være for å bli oppfattet som /i/? Vi kan gjøre et eksperiment. La oss gradvis redusere varigheten av [ i : ] i f.eks. 'fin'. (Vi kan klippe i lydbånd eller bruke elektronisk utstyr.) For hver gang spør vi forsøkspersoner om de oppfatter vokalen som /i:/ eller /i/. (Dette må i virkeligheten skje i tilfeldig rekkefølge og en bør spørre om ordet oppfattes som 'fin' eller 'finn'.) Dersom en fremstiller grafisk responsen som funksjon av stimulus (dvs. vokalvarighet), kan resultatet bli som i fig. 3. (Det er ikke alltid resultatet blir slik.) Det er nå naturlig å spørre hvor mange respons en trenger for hvert stimulus. Dessuten, med et nytt sett av forsøkspersoner, hvor mye må en regne med at vendepunktet (50 % respons) forandrer seg? Idet, som tidligere nevnt, varigheten er avhengig av konteksten, må denne holdes konstant (dvs. en må referere til ett ordpar). Resultatet vil også være avhengig av hvorledes vokalene realiseres i forsøkspersonenes dialekt.

### **Kontroll av treningsprogram.**

La oss anta at vi holder et kurs i stemmebruk for å forbedre stemmekvaliteten. Det vil være nærliggende å spørre om kurset har noen effekt. Først må vi prøve å definere begrepet stemmekvalitet, angi kriterier og fastlegge hvorledes det skal bedømmes. En kan tenke seg en subjektiv bedømmelse (av en gruppe personer som angir kvaliteten etter en eller annen skala) eller en objektiv bedømmelse (ut fra visse fysiologiske eller akustiske kriterier). Kurslederen eller terapeuten bør selvsagt ikke bedømme resultatene. Dersom noen 'pasienter' er motiviert og ønsker å forbedre stemmekvaliteten, kan det tenkes at det skjer uavhengig av treningsprogrammet. Dette fenomen kan studeres nærmere dersom en lar personer delta i to ulike treningsprogram der bare det ene skulle ha noen virkning på stemmekvaliteten. Ved å samle opplysninger om stemmekvalitet for en gruppe personer før og etter kurset, med og uten stimulering, kan en avgjøre om kurset har hatt noen effekt. En kan her benytte seg av testbatterier som f.eks. er utviklet, eller benyttes, for å teste effekten av medisiner.

### **Kurvetilpasning.**

I en rekke tilfeller bestemmer de sammenhørende verdier av to variabler og en er interessert i en eventuell funksjonsammenheng mellom variablene. Med eksempel fra fonetikk: Er det noen sammenheng, og i tilfelle hvilken, mellom

formantenes frekvens og båndbredde? La oss anta at en rekke sammenhørende målinger gir resultatet i fig. 4. Hvorledes kan en uttrykke sammenhengen mellom variablene? En vanlig metode går ut på å legge en kurve på 'best mulig' måte gjennom observasjonene. Dersom det er en teoretisk funksjonssammenheng eller en antatt sammenheng  $B = f(F)$ , kan konstantene i funksjonen beregnes og sammenhengen er dermed gitt. I mange tilfeller er det aktuelt å transformere en eller begge variablene slik at en oppnår en lineær sammenheng f.eks.  $y = a \log x$ . Den rette linje legges gjennom observasjonene ved hjelp av den såkalte minste kvadraters metode.

### **Sosiolingvistik.**

I sosiolingvistik undersøker en bruken av språklige variable i relasjon til sosiale faktorer. Et eksempel på dette er bruken av [ r ] ('tjukk l') sett i relasjon til brukernes partipolitiske holdning. Vi kan undersøke en gruppe personers bruk av [ r ] og [ l ] i ord der begge varianter kan forekomme og samtidig få rede på brukernes partipolitiske holdning, inndelt f.eks. i borgerlig og sosialistisk. Det knytter seg en rekke problemer til utvelgelsen av personer. Det er klart at flere andre sosiale faktorer som utdanning, boforhold og dessuten kjønn og alder kan være medbestemmende for språkbruken. Dessuten er vi sjelden konsekvente i språkbruken. For hver enkelt person blir det snakk om graden av bruk av den ene eller den andre språklyd. I den forbindelse kunne en også tenke seg en litt annerledes undersøkelse. Vi kunne f.eks. undersøke om det er noen korrelasjon mellom bruk av [ r ] og personenes inntekt. Statistisk sett vil dette være et helt annet problem. Ja vi kan gå enda videre og spørre hvilken av de sosiale faktorene som har størst betydning for bruken av [ r ]. Dette og lignende problem kan løses ved hjelp av multiple korrelasjonsanalyser eller faktoranalyse. Dette var eksempler fra noen forskjellige aktiviteter innenfor faget fonetikk, og jeg går ut fra at de fleste nå er blitt overbevist om at statistikk og databehandling, i en eller annen form, er høyst relevant for fonetikere.

# Databehandling og humaniora

*Jostein H. Hauge*

I de kommende numrene av Humanistiske Data vil vi gjennom artikler og intervjuer rette søkelyset mot de behov for utdanning i EDB som humanister i dag har. Vi vil her både se på behovene for frittstående studietilbud i emnet og for EDB-emner som utdanningselementer innenfor fagene.

Generelt må det kunne sies at vi i Norge har kommet relativt sent i gang med utdanning i databehandling. Mens utdanning i databehandling er under innføring på bred basis i ungdomsskole og gymnas f.eks. i Sverige — og er et hovedfag i Japan — tilbys emnet hos oss f.eks. for tiden bare som et valgfag i den videregående skole.

Som fagstudium på universiteter og høyskoler er imidlertid databehandlingsfaget (ofte kalt informatikk) styrket vesentlig i løpet av det siste 10-året. Utdanningskapasiteten er likevel her altfor lav til å tilfredsstille de aktuelle og forventede behov for arbeidskraft med slik utdanning.

Slik jeg ser det, er det 3 hovedgrunner til at humanistiske studenter og forskere fremover bør skaffe seg kunnskap i EDB.

Den første er innlysende nok: Datateknologi vil — også gjennom sammensmeltning med andre teknologier — i stigende grad sette sitt preg på samfunnsliv og dagligliv. Kunnskap om EDB og produkter hvor EDB inngår, vil i fremtiden være viktig for å kunne funksjonere i — og fremfor alt for å forstå — det samfunn vi lever i. Fremover bør det bli en stadig sterkere debatt om utformingen av samfunnet og om bruken av teknologi. For humanistenes deltakelse i den alminnelige samfunnsdebatt vil det være vesentlig at de kjenner virkemåten til og særtrekkene ved en av de bærende teknologier.

Dette fører naturlig over til neste moment: Kunnskap om datateknologi vil utgjøre en viktig bakgrunnskunnskap også for lektorer i HF-fag i f.eks. den videregående skole. Jeg tenker her bl.a. på at datateknologien vil få en sentral plass i samfunnsrelaterte undervisningsemner. I parentes bemerket vil også generell teknologihistorie bli verdifull i denne sammenheng. Undervisningskompetanse i EDB-faget vil kunne utvide HF-lektorenes undervisningsplattform i skolen på en gunstig måte. Det er dessuten å håpe at lektorer med fagutdanning i HF-fag i særlig grad vil kunne sette datafaget inn i et samfunnsmessig og humanistisk perspektiv.

Det tredje argumentet for HF-utdanning i EDB har med bruk av databehandling i forskning å gjøre. I tiden som kommer vil datamaskiner bli et «naturalisert» arbeidsverktøy også for humanistene. Bruken av EDB vil øke i fortsettelse av den virksomhet som har skjedd til nå, og det vil fremover foreligge stadig bedre tilpasset maskin- og programvare til bruk i HF-fagene. Opplæring i bruk av EDB vil gå parallelt med innføring i de sentrale metodene for databehandlingsarbeid, herunder kvantitativ bearbeiding av data.

Dersom det kommer i gang organisert forskerutdanning i de humanistiske fag, vil det være riktig å gi EDB-emner en klar plass i utdanningstilbudet — og i enkelte sammenhenger en særlig sentral plass. Det vil være vesentlig for en fornuftig og formålstjenlig bruk av EDB i de humanistiske fag at den videre metodeutvikling innenfor fagene kan ledes av humanistene selv med basis i fagenes egenart, målsetting og tradisjoner. Siden enkelte av de humanistiske fag bl.a. forvalter sentral viten om språklig form og funksjon bør en også understreke de forpliktelser disse fagene har til å gå med i den generelle utvikling av EDB-baserte språklige hjelpemiddel. At språkfolk bør arbeide med spørsmål knyttet til norsk dataspråk er klart nok, men interessen burde også strekke seg videre til samarbeid med teknologiske miljøer om brukervennlige kommunikasjonsformer med datamaskin.

I fremtiden vil stadig mer sentralt forskningstilfang være lagret på EDB-tilgjengelige medier (tekster, arkiver, samfunnsdata). Det samme gjelder sekundærdata i form av bibliografiske databaser og andre vitenkilder. For utnyttelse av slike kilder vil kunnskap om EDB være en fordel og i mange tilfeller en forutsetning. De humanistiske fag som er orientert mot studiet av kulturformer i samtiden, vil i fremtiden også måtte trekke konsekvensene av data-teknologi inn som et studieemne. Også her vil kunnskap om teknologiens form og funksjon danne basis for forståelse.

Mens datateknologi frem til i dag i særlig grad er blitt tatt i bruk i næringsliv og industri, vil EDB-anvendelsene i offentlig virksomhet — inkludert forskningssektoren — fremover bli systematisk bygd ut. Det kan bl.a. her vises til at myndighetene for tiden planlegger en større produktivitetskampanje rettet mot offentlige etater og institusjoner.

For de humanistiske fag synes faren for massearbeidsløshet på grunn av bruk av EDB å være relativt lite sannsynlig. Heller vil bruk av EDB innenfor denne sektor kunne føre til at en langt større arbeidsmengde enn tidligere kan utføres med samme stab. Siden de humanistiske fag ikke kan regne med stor stillingsutbygging i fremtiden, vil mye av den faglige ekspansjon innenfor fagene måtte skje ved hjelp av bedret arbeidsverktøy. Innenfor bank- og forsikrings-sektoren har arbeidsmengden øket med ca. 50% fra 1970 mens antallet sysselsatte i disse næringer bare har steget meget lite. En slik «effektivisering» av humanistisk forskning vil neppe finne sted ettersom de fleste humanistiske fagene ikke primært er tjeneste- og produksjonsytende virksomhet. Men det er klart at det f.eks. innenfor de deler av humanistisk forskning hvor arbeidet består i å bygge opp, endre og utnytte svære datamengder, vil EDB-metoder i fremtiden kunne bli et hjelpemiddel av betydelige dimensjoner.

Vi håper i løpet av de kommende presentasjoner å vise hvordan databehandling er og vil bli tatt i bruk innenfor undervisning og forskning i de humanistiske fag. Vi vil også prøve å drøfte hvilke farer økt bruk av EDB innen humaniora kan representere.

Dersom noen har synspunkter på noe av det som kommer fram i intervjuene, åpner vi gjerne våre spalter for innlegg.



*Ivar Fønnes, amanuensis ved EDB-tjenesten  
ved Det historiske-filosofiske fakultet,  
Universitetet i Oslo.*

---

# EDB OG HUMANIORA

---

*Intervju: Svein Skotheim.*

*Som nevnt av Jostein H. Hauge i introduksjonen på de foregående sider starter vi i dette nummer av Humanistiske Data en intervjuserie som vi har valgt å kalle «EDB og humaniora». Gjennom samtaler med sentrale personer vil vi rette søkelyset mot behov for og utforming av nye utdanningstilbud og tilrettelegging av forholdene for en fornuftig bruk av edb innen humanistisk forskning.*

*Først ut er amanuensis Ivar Fønnes ved HF-data, Universitetet i Oslo. De startet nå i høst sitt utdanningstilbud «EDB for humanister» som et semesteremne, dvs. at emnet tilsvarende ett semesters studium. Det er imidlertid lagt opp til at arbeidet fordeles over ett år. Studieenheten kan utgjøre en selvstendig del av cand.mag-studiet, og som alle semesteremner forutsettes det at en har avlagt minst en fageksamen på forhånd. Det er laget en fyldig og oversiktlig studieplan som kan fåes ved henvendelse til EDB-tjenesten ved Det historisk-filosofiske fakultet, Postboks 1102, Blindern, Oslo 3.*

---

*Hva er bakgrunnen for at dette studietilbudet ble opprettet?*

Det er særlig to forhold som har aktualisert et slikt studietilbud.

Først og fremst har det økende antall hovedfagstudenter og forskere innen humaniora som bruker EDB i sitt arbeid, skapt behov for et opplæringstilbud fra vår side. Dessuten er det nå blitt mulig ved Universitetet i Oslo å tilby fagenheter på ett semesters lengde, såkalte semesteremner. Vi har tidligere følt at et grunnfag ville bli for stort med den kapasitet EDB-tjenesten i dag har.

Trolig ville det også bli for pretensiøst overfor de brukergrupper som er aktuelle. Det studietilbudet vi nå introduserer, vil være et semesteremne som strekker seg over ett år, og som derfor etter planen skal leses ved siden av et annet fagstudium.

Til nå har en del studenter fra HF-fagene tatt innføringsemner ved Institutt for informatikk, men der tar man sikte på en fagutdannelse i EDB. Vi vil i sterkere grad se på EDB som et redskap som i bestemte sammenhenger kan brukes innen de humanistiske fagområder. På sikt kan det kanskje bli aktuelt å samarbeide med Institutt for informatikk om grunnopplæringen.

Ellers er det grunn til å anta at studietilbudet generelt kan bidra til å styrke filologenes stilling på arbeidsmarkedet. EDB er jo i dag et almenfag som alle bør ha et elementært kjennskap til.

Det er også av stor betydning å trekke problemstillingene omkring datateknologiens samfunnsmessige konsekvenser inn i de humanistiske fag. Flere av disse fagene tar for seg kultur- og samfunnsspørsmål i ulike sammenhenger, og det er derfor viktig at konsekvensrike fenomener som EDB tas opp i fagene. Dette forutsetter igjen kjennskap til teknikken og dens anvendelsesområder.

*Vil studietilbudet være rettet mot de humanistiske faggrupper som har særlig nytte av EDB?*

Til nå har historikere, språk-, litteratur- og musikkforskere vært de mest aktive, men innen filosofi, arkeologi, etnologi, navnegransking, religionshistorie, kristendomskunnskap og andre fag er det enkelte som etter hvert bruker EDB i sin forskning. Det er klart at eksempelmateriale og prosjektbeskrivelser vil være preget av de fag som er kommet lengst i anvendelsen av denne teknikken, som f.eks. historie og språkvitenskap. Vi vil imidlertid prøve å tilpasse opplegget til de aktuelle studenter, særlig m.h.t. oppgaveformuleringer og eksempelmateriale.

*Er studietilbudet først og fremst for forskere, eller vil det også være av interesse for folk som planlegger en lærerutdanning?*

Som nevnt må EDB i dag betraktes som et almenfag i skolen, og våre kandidater bør ikke stå uten kunnskaper i et slikt fag. Vi vet at norskfaget ofte fungerer som et debattfag, og bruk av EDB er et aktuelt debatttema i mange sammenhenger. Det er imidlertid klart at dette fagtilbudet ikke vil gi undervisningskompetanse i skolen. Det er også klart at utgangspunktet for opprettelsen av fagtilbudet er behovene innen forskningsrettet arbeid.

Ellers er bruk av EDB aktuell innen forlag, presse og på andre områder der filologene tradisjonelt har sitt arbeidsmarked.

*Legger studiet vekt på kritisk drøfting av de farer bruk av EDB kan representere, f.eks. faren for styring av forskningen mot mål som gjør bruk av EDB naturlig?*

Det er ikke lagt opp til en systematisk gjennomgåelse av slike problemstillinger. Disse spørsmålene kommer først og fremst inn ved presentasjon og

drøfting av enkeltprosjekter. Gjennomgåing av prosjekter står sentralt i siste del av emnestudiet, og da vil det bli lagt vekt på drøfting av metoder, problemformulering osv.

Ellers understreker vi sterkt nødvendigheten av en nøktern holdning til bruk av EDB. Denne teknikken er bare et redskap på linje med mange andre tekniske hjelpemidler som kan gjøre forskningsarbeidet mer effektivt.

*Hvordan arbeider EDB-tjenesten ved Det historisk-filosofiske fakultet ved Universitetet i Oslo? Har dere f.eks. planer om andre studietilbud?*

Vår oppgave er å gi metodisk og EDB-messig veiledning for forskere, programmeringsassistanse, kurs- og konferansetilbud m.m. I den utstrekning ressursene strekker til, arbeider vi også med metode- og programutvikling.

For tiden er det 2 stillinger ved EDB-tjenesten etter at det fra 1. august er tilsatt en vitenskapelig assistent.

Vi har for tiden ingen planer om nye studietilbud. Dette semesteremnet er det maksimale vi kan tilby ut fra tilgjengelige ressurser. Den ordinære veiledningstjeneste overfor forskere vil etter planen fortsette omtrent som før.

---

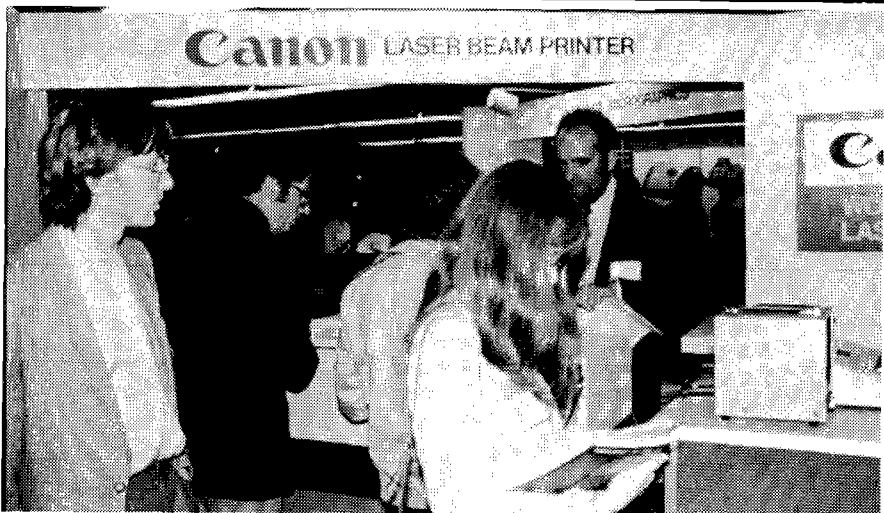


# RAPPORTER

## **National Computer Conference (NCC), 4-7 mai 1981, Chicago.**

*Knut Hofland*

Dette er den største amerikanske edb-kongress og utstilling av EDB-utstyr med over 600 utstillere og foredrag i 8 parallelle sesjoner. En egen sesjon var viet «Visuals, Natural Language Processing and Artificial Intelligence». I tillegg var det en «Personal Computer Festival» der bruk av (små) mikrodatamaskiner ble presentert. Jeg besøkte utstillingen en dag og fulgte også en sesjon om kommunikasjon med datamaskin ved naturlig språk. Flere firma hadde utstilt laserprintere, bl.a. store printere fra Siemens, HewlettPackard, IBM og Xerox og mindre utgaver, bl.a. en fra Canon som var på størrelse med en vanlig kopimaskin. Prisen på denne siste var \$16.000 og kunne skrive 10 sider pr minutt med forskjellige skrifttyper. Denne vil gi en kvalitet som er bedre enn tekstbehandlings-skriver og kan i mange sammenhenger brukes i stedet for fotosats. Den er f.eks rimeligere i drift enn en fotosetter.



*Laserprintere sto sentralt på konferansen i Chicago. Canon demonstrerte en utgave på størrelse med en kopimaskin.*

Dataskjermene blir stadig bedre både med hensyn til oppløselighet og bruk av farge, slik at det i tillegg til tekst kan vises bilder og figurer. Flere firma hadde også kopienheter til skjerm basert på videobilde. Dette sammen med lagring av informasjon på laser-plater (som også ble vist på utstillingen), åpner store muligheter for behandling av billedinformasjon. Det er også aktuelt å lagre tekstdokumenter som f.eks. bilder på denne måten. Det kan da bare søkes etter og sorteres i felter som dataregistreres på vanlig måte, men i tillegg kan hele dokumentet fås frem som et bilde. Det arbeides også med overføring av dokumentbilde til alfanumerisk koding (en utvidelse av tradisjonell OCR-lesing). Det ble ellers vist enheter for innmating og utmating av tale. Utstyret for innmating har foreløpig begrenset kapasitet, ca. 100 ord. Et ferdig system med terminal og mikromaskin kostet ca. \$5500.

---

## **The Fifth International Conference on Computers in the Humanities (ICCH/5).**

*Knut Hofland og Øystein Reigem*

ICCH/5 ble holdt i Ann Arbor i USA 17-19. mai 1981. Det ble presentert 50 foredrag i løpet av 2 dager. Som vanlig var spredningen stor både når det gjaldt emneområder og kvalitet på foredragene. Det ble ikke delt ut noe skriftlig materiale på forhånd, hverken «preprints» eller deltakerliste, men «proceedings» ble annonsert.

Under følger korte sammendrag fra en del av foredragene.

### **Character definition through syntax.**

*Rosanne Potter (Iowa State University)*

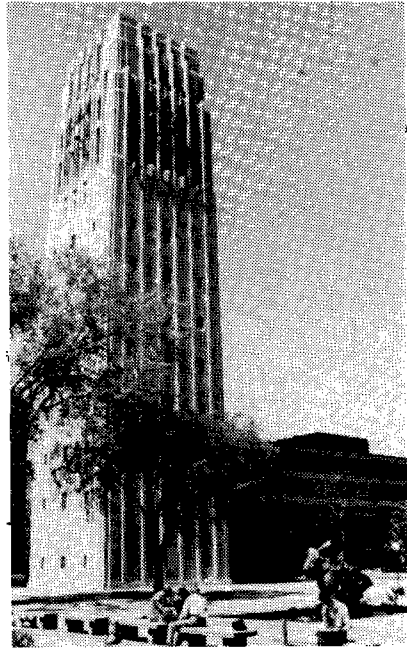
Foredragsholderen hadde studert 21 skuespill av amerikanske, engelske og irske forfattere fra vårt århundre. Hun hadde talt opp bruk av en del forhold som utgjorde hennes variable, bl.a. personenes bruk av adverb, sammenligninger, definisjoner, utbrudd, hypoteser, imperativ, spørsmål o.l. Hun regnet ut gjennomsnitt og standardavvik for et skuespill og totalt for alle skuespillene. Videre ble det regnet ut F-verdier som viste hvilke variable som skiller de enkelte skuespill, og korrelasjoner som skiller de enkelte personer. Hun regnet også ut z-verdier for hvert stykke og hver person. Ved hjelp av histogrammer og kurver ble de enkelte personers bruk av de forskjellige variable beskrevet. Prosjektet er nærmere omtalt i *Computers in Humanities nr. 3 vol. 14* og i en kommende artikkel i *Style*.

### **Database design for biblical texts.**

*H. van Parunak (University of Michigan)*

Foredraget startet med å beskrive en del karakteristiske trekk ved analyse av bibeltekster, som f.eks. at det fins en sterk tradisjon mht. morfologisk, syntaktisk og semantisk analyse, at tekstene har et rikt og variert opphav og at de fins i mange oversettelser og utgjør et stratifisert system. Bibeltekster fins i flere ortografiske nivåer både på hebraisk og gresk, fra bare konsonanter til tekster der aksenter og tegnsetting er inkludert.

Parunak gikk så over til å beskrive relasjonsdatabasemodellen og satte opp relasjoner for å beskrive grafiske ord, leksikalske ord og en type ordforbindelser. Systemet var ikke implementert ennå pga. mangel på relasjonsdatabasesystem. I diskusjonen ble det imidlertid nevnt flere som kunne være aktuelle både på stormaskiner og også mikromaskiner, selv om en del operasjoner her ville ta lang tid.



*Dette tårnet ble et velkjent landemerke for deltakerne på ICCH-konferansen i Ann Arbor, USA.*

---

### **Computer-Based Techniques for Understanding Performed Music.**

*Martin Piszczalski og Bernard A. Galler (University of Michigan)*

Foredraget omhandlet et prosjekt hvor lyden ble overført til et tre-dimensjonalt bilde på en fargeskjerm der høyden varierte etter intensiteten, og frekvensområdet gikk fra venstre mot høyre. Bildet ble filmet synkront med musikken slik at det beveget seg, det artet seg som om en fløy over et fjellandskap. En av foredragsholderens påstander var at lyden varierer mer enn en kan få inntrykk av ved å lytte. På denne måten oppnås en slags visuell musikkteori som kan gi opphav til nye forskningsområder.

### **CACTUS Computer Assisted Typesetting Use System.**

*David R. Chesnutt (University of South Carolina)*

Nøkkelen til systemet var filstrukturen og ved hjelp av denne ble det satt inn 85% av de typografiske koder. Det var program til å sette på kode til forskjellige fotosettere, og en hadde prøvet ca. 10. De fleste tok data på magnetbånd, men noen tok også fleksiplate. Det ble brukt både mikromaskiner og OCR ved innmating av tekst. Systemet ble brukt til produksjon av «Papers of Henny Laurens» (Robert Oakman). Programmene var i PL1 og ble kjørt på Amdahl V6. Disse vil bli tilgjengelig en gang i 1982.

### **Mikrodatamaskiner**

Fra Brigham Young University var det 3 foredrag om bruk av mikrodatamaskiner innenfor humanistiske fag. *Randall Jones* satte opp en del krav som var stilt ved anskaffelse, bl.a. enbrukersystem, fargeskjerm, grafisk mulighet, tilbud av programmer, utbyggbarhet og god forhandlerstøtte. Valget falt på en Apple (pris fra

1500 - 2500 dollar), og det fantes 10 slike i de humanistiske miljøer. Foredragsholderen kom også inn på de begrensninger som i øyeblikket gjelder mht. hukommelse, hastighet, programmeringsspråk, oppløselighet på skjerm o.l. Fordelene med bruk av mikromaskiner er at de er billige, portable, uavhengige av annet utstyr og fleksible for den enkelte bruker. Maskinene ble brukt innen 6 områder: tekstbehandling, administrasjon, dataregistrering, databehandling, datamaskinassistert læring (DAL) og kontroll av rettskriving osv.

*John S. Robertson* orienterte om et prosjekt for å lage en ordbok over indianske dialekter. De hadde tatt med en Apple mikrodatamaskin til Guatemala i 2 måneder for å registrere data. Det hadde vært en del problemer med å overføre disse til et IBM-anlegg. Konklusjonen var bl.a. at en burde brukt et annet tekstbehandlingssystem og også sørget for batteri ved strømbrudd! *Kim Smith* holdt et foredrag om bruken av mikromaskin i datamaskinassistert læring av fremmedspråk. Han startet med å konstatere at datamaskinassistert læring ikke har fått den utbredelse som en hadde ventet. Det gjelder å ta datamaskinassistert læring i bruk på de områder der dette har noe for seg. Apple har farger og lyd og dette kan utnyttes i instruksjonen. Men det er vanskelig f.eks. å kontrollere en students uttale av et ord mot en fasit. For visse språk er det problem med inndata, f.eks. kinesisk. Til dette brukes en egen maskin med digitaliseringsbord. DAL brukes i øyeblikket for kinesisk, japansk og tysk. En arbeider med å kunne få bilder på videoskjerm som kan vises på kommando fra bruker eller program. Videre var en interessert i å tilknytte en audiodisk som også kan styres via terminal.

### **The Kurzweil Data Entry Machine: Some Experiences of Optical Character Recognition**

*Susan Hockey (Oxford University)*

Hockey redegjorde for sine gode erfaringer med Kurzweils OCR-leser. The Kurzweil Data Entry Machine (KDEM) består av en Nova minimaskin, en Hewlett-Packard skjerm, en lese-enhet (scanner) og et platelager for lagring eller mellomlagring før overføring til magnetband. Det implementerte systemet er et multifont-system som leser 6-24 punkters skrift med en fart på 50.000 tegn pr. time ved god tekst (25-30 sider). Farten kan forbedres ved utskifting til raskere prosessor eller med en ekstra terminal.

Maskinen er papir-tolerant og kan lese svært varierende skrifttyper såsant tegnene ikke henger sammen (f.eks. arabisk). Fotokopier tar den imidlertid bare når kopien er god. Programvaren inneholder noen få feil og egenheter, men det ser ut til at disse blir rettet etterhvert. Også lesingen forbedres etterhvert.

### **Mathematical Typesetting for the Humanities**

*Lynn A. Price (TEX User's Group)*

Price fortalte om de to programsystemene TEX (uttales TEK av en eller annen grunn) og METAFONT. Mathematical typesetting kunne tenkes å bety setting av matematisk materiale, noe som burde passe bra for de to omtalte systemene, men i dette tilfellet refererte termen til matematikken bak systemene.

TEX er et programsystem med et interessant konsept: «bokser» og «lim» som tekstbildet bygges opp av. En «boks» kan inneholde et tegn eller flere «bokser» og «lim». «Limet» er det hvite mellom tegnene og er tøylig, slik at f.eks. en remse bokser med ord og lim mellom kan strekkes over linjen for å få rett høyremarg og lik avstand mellom ordene.

METAFONT er et programmeringsspråk for font-design. Systemet tar imot parametriserte tegnspesifikasjoner og bruker en tenkt penn med forskjellige spisser (sirkelformet, elliptisk, kvadratisk) til å skrive og viske ut. Output er en såkalt DVI-file (device independent file) med koordinater for de enkelte karakterene. Systemet er skrevet i et maskinavhengig språk, men kommer i en Pascal-versjon i sommer for DEC 10, 20 og Univac 1100. Denne skal være anslagsvis 80% maskinuavhengig. (Dessverre har ikke referenten tilsvarende opplysninger for TEX.)

### **OMNIFRONT: A General-Purpose Font Storage and Output Package**

*John L. Dawson, Yeshe Zangmo and Suchitra Sengupta (University of Cambridge)*

OMNIFONT er beregnet for strek-karakterer tegnet med plotter. Linjestykkene går mellom punkter i en 16x16 matrise. En kan også definere kompakte tegn satt sammen av f.eks. tettstilte, vannrette linjestykker, men dette er mye arbeid. Det kan skaleres uavhengig horisontalt og vertikalt. Skrivningen kan orienteres i 4 retninger (rotasjon 0°, 90°, 180° og 270°), og plasseringen av et tegn spesifiseres absolutt eller relativt av forrige tegn. Input til systemet er en streng som kan være en blanding av koder for predefinerte tegn, egne tegneinstruksjoner og koder for posisjonering og orientering. En har i OMNIFONT bl.a. definert et subsett av det kinesiske alfabetet.

### **Revolt in Flatland: An Opera in Two Dimensions**

*Sheldon Klein, David A. Ross, Mark S. Manasse, Johanna Danos, Mark S. Bickford, Walter A. Burt og Kendall L. Jensen (University of Wisconsin)*

I forrige århundre skrev Edwin Abbott «Flatland», en fortelling om livet i en 2-dimensjonal føydal verden bebodd av linjestykker, manglekanter og sirkler. «Revolt in Flatland» er en datamaskinsimulering av et noe bearbejdet Flatland. De fysiske lover og den sosiale strukturen i Flatland er simulert i et nytt språk, «l l l l » (for korthets skyld «BAR BAR»), interpretert i UCSD Pascal og kjørt på Terak og Apple mikromaskiner. Språket har en simuleringsdel og en semantisk nettverk-del. Begivenheter skapt av simuleringsdelen kodes som endringer i nettverket. Den varierende nettverksstrukturen brukes til å generere løpende bilde, tekst og musikk. Forskjellige kjøringer gir gjerne varianter av hovedintrigen. (For skeptiske lesere kan referenten opplyse at prosjektet har faste røtter i forskning om kunstig intelligens.)

Grunnet diverse tekniske vanskeligheter ble selve forestillingen redusert til musikk og stillbilder, men foredragsholderen fortalte med innlevelse om intriger, sub-intriger, mål, betingelser, demoner og stemninger.

### **Suprasegmental Graphemes as a Technique for Computer Encoding of Literary Texts**

*Earl M. Herrick (Texas A & I University)*

Herrick hadde gjort den opplagte observasjon at tekster ofte inneholder flere tegn og skrifttyper enn datamaskinen umiddelbart kan takle. Alternativene blir da enten en forenkling av materialet eller en merking etter et eller annet system. I foredraget sitt presenterte han sin egen metode for merking av litterære tekster. Merkingen gikk stort sett på tegnbruken, såsom 1) avvikende skrifttyper, 2) diakritiske tegn og 3) spesielle tegn. Til merkingene 1) og 2) ble det brukt koder bak ordet eller tegnet (postfiks), og en merket tekst ble dermed forholdsvis leselig, noe som ellers ikke alltid er tilfelle. I en sekvens med avvikende skrifttype måtte imidlertid hvert ord merkes. Foreleseren påpekte ellers behovet for en konvensjon for tekstmerking.

Det kan til slutt nevnes at Knut Hofland holdt foredrag om Ibsen-konkordansprosjektet: A Lemmatized Ibsen Concordance.

## «Joint Conference on Easier and More Productive Use of Computing Systems».

*Knut Hofland og Øystein Reigem.*

Denne konferansen var en sammenslåing av «Data Bases in the Humanities and Social Sciences» og «Systems and their Uses». Konferansen ble arrangert i Ann Arbor i fortsettelsen av ICCH/5 fra 20-22. mai, med over 80 foredrag i løpet av 3 dager. Den utvidete rammen gjorde at denne fikk et mer teknisk preg. Det var representanter fra flere maskinleverandører som hadde foredrag, bl.a. IBM, UNIVAC, Xerox, Texas CDC og Hewlett-Packard. En del av foredragene handlet om programvare for enklere utvikling av systemer og bedre mulighet for hjelp og feilfinning, spesielt for uvante brukere. Det forelå «abstracts» til foredragene.

Fra Norge holdt *Knut Hofland* et foredrag om tekstsøkesystemer brukt i humanistiske fag (medforfatter *Sigbjørn Århus*), *Øystein Reigem* redegjorde for SIFT-prosjektet og *Jan Oldervoll* for statistisk behandling av 1801-folketellingen.

Flere av foredragene omhandlet lagring av opplysninger om billedmateriale, spesielt om oppbygging av tesauri. Foredragsholderne tilhørte arbeidsgruppen «Image Access Society» som også hadde møte under konferansen.

*Denis Castonguay* (Public Archives of Canada) presenterte en tesaurus for kanadisk ikonografi. Denne skal brukes som en terminologisk kontroll i et EDB-basert arkivsystem. I Picture Division har en ca. 100 000 arbeider og til hvert av disse er det opptil 80 dataelementer.

*Elanor Fink* (National Collection of Fine Arts, Smithsonian Institution) orienterte om en tesaurus for amerikanske kunstverk. Denne inneholder hovedkategorier og underkategorier. Innenfor hovedkategoriene er det listet opp hvorledes underkategoriene skal brukes. En har tilgjengelig et informasjonssystem som kan bruke verbale deskriptorer og søke på kombinasjoner av disse.

*Anne-Marie Logan* (Yale Center for British Art) orienterte også om en tesaurus. Hennes materiale dekker tidsrommet 1500-1915 og alle opplysningene finnes i naturlig språk. Hun hadde tidligere brukt satsvis kjøring, men var på vei over til interaktiv. Tesaurusen bestod av 300 sider. Termer som kunne brukes var enten primære, sekundære eller tertiære. De to første var lukkede klasser. Hennes tesaurus var tilgjengelig mot et mindre beløp.

*Christopher Seifried* (Public Archives of Canada) hadde som foredragstittel «Guide to Canadian Photographic Archives, A National Inventory Project». Han åpnet med å påpeke at dersom en har en tesaurus og et databasert gjenfinningsystem, så er nytten av disse større enn nytten av hvert enkelt. Konseptkontroll er viktig, men en trenger systemene nå og kan derfor ikke vente på en perfekt tesaurus før en starter. Fra starten var prosjektet satsvis orientert med en fotosatt katalog som mål. I overgangen til et interaktivt system valgte man Library of Congress emnebibliotek, modifisert av bilde- og fotografiavdelingen der. På denne måten ville to av de største indekser over bilder i Nord-Amerika følge samme standard. Emnelistene måtte imidlertid justeres noe pga. kulturelle og historiske forhold og fordi Canada har to offisielle språk. Person- og stedsnavn er inkludert fra to kanadiske indekser over slike. En forbedret framfinning kan skje når synonymlister kan bygges opp automatisk. Videre ble en artikkel i «Visual Resources» nevnt: «Video Scan Picture Searching» som åpner nye veier for katalogisering og søking.



*Kommunikasjon med maskin i naturlig språk vies stor oppmerksomhet. Knut Hofland fra Senteret tester et forsøksopplegg på konferansen i Ann Arbor.*

---

### **Concise Natural Language Interaction**

*Paul R. Michaelis og James A. Hendler (Texas Instruments)*

Kommunikasjon med maskin i naturlig språk blir viet en del oppmerksomhet for tiden. Foredragsholderne mente det var ting som tydet på at et semantisk basert interaktivt NL (natural language) system ville tjene på restriksjoner. De bygget bl.a. på en variant av det velkjente eksperiment der to personer i hvert sitt rom har henholdsvis byggesett og bruksanvisning. Kommunikasjonen mellom de to gikk via teletype i noen forsøk og stemme i noen. Det spesielle med eksperimentet var at halvparten av personene ble bedt om å være så eksakte og enkle som mulig. Dette så ikke ut til å hemme løsningen av problemet.

### **Learning Effectiveness: The Impact of Response Time**

*Sherry Weinberg (Control Data Corporation)*

Weinberg tok opp responstid - et aktuelt emne for de fleste ved EDB-senteret. Hun hadde gjort en studie av responstidens virkning på læreprosessen. Forsøkene ble gjort vha. læresystemet PLATO, tilknyttet via to nettverk med forskjellig responstid (henholdsvis 0,25 og 1,3 sekunder i gjennomsnitt). Det viste seg at prestasjonene og holdningen til systemet ble signifikant bedre ved den korteste responstiden.

### **What Can Be Learned From Arcade Games and Home Computer Applications?**

*Karl Zinn (University of Michigan), Ben Schneiderman (University of Maryland) og Thomas Malone (Xerox Palo Alto Research Center), Paneldiskusjon.*

I sitt innlegg «The Case for Considering Games and Home Applications» gjennomgikk

Zinn diverse TV-spill med stor entusiasme. Han viste til TV-spillenes store popularitet og utbredelse. Han mente denne suksessen i hvert fall delvis skyldtes følgende egenskaper ved spillene: lett å komme i gang, høy grad av visualisering for betjeningsorganer og resultater, god respons.

Schneidermann holdt deretter et innlegg med tittelen «Direct Manipulation: A Step Beyond Programming Languages.» Begrepet «direct manipulation» illustreres best ved noen eksempler: Å kunne flytte lysflekk («cursor») på dataskjermen vha. en spake er direkte manipulasjon. Å kunne velge fra en meny på skjermen, og å kunne flytte tegnninger eller modeller av gjenstander med en lyspenn på en grafisk skjerm, er også direkte manipulasjon. En skjermeditor, der en til enhver tid har et utsnitt av data i sin riktige form på skjermen, er mer «direkte» enn en linjeeditor. Ved direkte manipulasjon benytter en seg av tre beslektede teknikker: 1) Gi en fysisk direkte måte å flytte lysflekk eller manipulere objekter av interesse. 2) Gi en konkret visuell representasjon av objektene, for så umiddelbart å skifte synspunkt til operasjonene som kan utføres på eller med objektene. 3) Unngå å bruke kommandospråk og heller satse på operasjonene knyttet til (den kognitive) modellen vist på skjermen. Ved direkte manipulasjon får nybegynnere tilgang til maskin- og programressurser uten å måtte lære komplisert syntaks og mengder av kommandoer.

I det siste innlegget, «What makes computer games fun?», redegjorde Malone for akkurat dette, og for hvordan de attraktive trekkene kunne brukes i interaktive programsystemer. Malone hadde gjort 3 studier av barn og TV-spill. Han fant ut at tiltrekkende egenskaper ved spillene var målrettethet, poengtelling, lydeffekter, innlagte tilfeldigheter, responstid, visuelle effekter, konkurransemomentet, ... rangert i denne rekkefølgen, med målrettethet som spesielt viktig. Ut fra undersøkelsen fremla han endel viktige kriterier for systemdesign, gruppert under nøkkelordene utfordring, fantasi og nysgjerrighet. Han påpekte at en selvfølgelig måtte være oppmerksom på skillet mellom verktøy og leketøy.

---

## **«On the Possibilities and Limits of the Computer in producing and publishing Dictionaries», Pisa 20-22 May, 1981.**

*Jostein H. Hauge*

Arbeidsseminaret ble arrangert av European Science Foundation (ESF), Standing Committee for the Humanities. Bakgrunnen for dette seminaret var at en rekke av medlemsorganisasjonene i ESF (dvs. de nasjonale forskningsråd) som støtter ordboksprosjekter, melder om stadig stigende vanskeligheter med å finansiere slike tiltak. Samtidig er leksikografi et felt hvor de nye teknikker for databehandling og publisering i økende grad kan få direkte anvendelse i prosjektarbeidet.

Gjennom forberedelsen til seminaret, hvor prof. *Egil Pettersen* deltok fra norsk side, kom man frem til at bare ordboksarbeid som faller inn under de humanistiske fag, skulle inkluderes. Det vil si at problemer og muligheter knyttet til arbeid med tekniske og vitenskapelige ordbøker ikke ble tatt opp i denne sammenheng.

Det deltok ca. 50 deltakere fra 15 europeiske land og fra USA (forelesere). Deltakerne var utpekt av de forskjellige forskningsorganer som er representert i ESF.

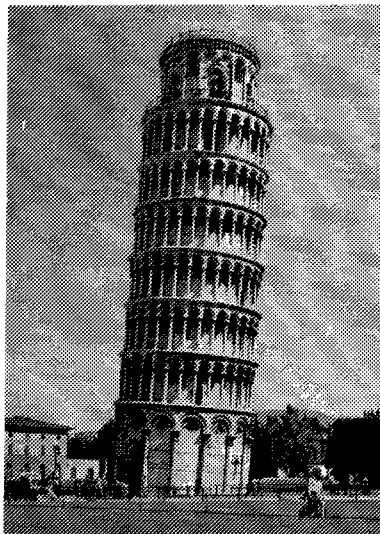


---

*The tower of Pisa is a beautiful symbol for the fact that human beings cannot foresee the social implication of their works. The artist did, of course, not foresee that the weakness of the foundation of the building would bring about the lopsidedness of the tower and by that fact would attract the attention of all humanity. Is this not true also for more abstract creations of man in the sense that their factual social consequences conform only to a little extent with the intentions of the creator?*

*Albert Einstein, 1953*

---



### *Spørreundersøkelse*

Som ledd i det forberedende arbeidet med seminaret, foretok Standing Committee for the Humanities en spørreundersøkelse blant sine medlemsorganisasjoner om de større ordboksaktiviteter som de støttet. Til denne undersøkelsen kom det inn svar fra 14 medlemsland som totalt viste at det i dag er i gang arbeid med minst 84 store ordbokstiltak i de europeiske land, støttet av det offentlige. (Når det gjelder Tyskland, se nedenfor). Det materialet som kom inn, ble videre analysert før seminaret (*Isabelle Warnesson*), og det foreligger en rekke interessante opplysninger i den rapporten som ble utarbeidet. *Rapporten kan fåes ved henvendelse til Senteret*. Det skal her bare nevnes at følgende 6 land har meldt om flest arbeider i gang: Italia (21), Frankrike (10), Sverige (9), Norge (8), Danmark (8), Spania (7). Tyskland synes imidlertid å ha falt ut i rapporten. Fra andre kilder vites at det er minst 30 prosjekter i gang støttet av det tyske forskningsråd.

Oversynet viser at ca. 55% av alle tiltakene gjør bruk av datamaskinelle metoder i en eller annen forstand (i mange tilfeller avgrenset til publiseringsfasen, f.eks. fotosetting). Det må imidlertid legges til at tallene ovenfor ikke gir et riktig grunnlag for sammenligning av innsatsen landene imellom og den totale aktivitet i et land. Flere påpekte under konferansen at det i mange land foregår viktig ordboksarbeid i kommersiell regi (f.eks. Oxford English Dictionary). Det er også store forskjeller i størrelse, fra prosjekter med et par ansatte til f.eks. prosjektet *Trésor de La langue Française* som for tiden beskjeftiger 145 heltidsansatte ved Universitetet i Nancy.

### *Fra innledningsforedragene*

I et av innledningsforedragene med tittelen Lexicography and Linguistics tok prof. *John Lyons* (England) for seg det lingvistiske grunnlag for leksikografisk arbeid, og foretok en grenseoppgang mellom leksikologi og leksikografi. Leksikologi fokuserer de teoretiske sider ved ordboksarbeid og legger vekt på språkets leksikalske struktur, mens en i leksikografien har hovedinteressen vendt mot oppbygging og produksjon av ordbøker.

Foredragsholderen mente imidlertid at det bak alle prosjekter «lurer» en eller annen lingvistisk metode eller teori. Det er viktig at man ikke ureflektert anvender den for tiden mest «populære» lingvistiske metode i sitt leksikografiske arbeid. En hoveddel av foredraget var for øvrig viet analyse av begrepene leksikalsk og semantisk struktur og forskjellige måter å betrakte synonymi på.

Lyons kom mot slutten av sin presentasjon inn på forholdet mellom språkvitenskap og området kunstig intelligens. Det ble her bl.a. sagt at for språkvitenskapen har denne nye forskningsgren hatt betydning bl.a. ved å gjøre det klart hvor vanskelig det er å formalisere lingvistiske strukturer.

Prof. *Paul Bratley* (Canada) kom i foredraget «Methodological and Technological Problems in computerized Lexicography» inn på de spesielle krav som leksikografisk arbeid stiller i forbindelse med databehandling. Av spesielle tema som ble tatt opp, kan nevnes størrelsen på de leksikografiske databaser, som gjør det nødvendig med effektive redskap for gjenfinning av informasjon, de varierte og store tegnsett som forutsettes, kravene til utskriftskvalitet, behov for behandlingsverktøy for editering, sortering, automatisk analyse og klassifikasjon av ordmateriale (bl.a. ved grammatisk merking og lemmatisering).

Leksikografisk databehandling er en spesialitet, sa Bratley og advarte leksikografene mot selv å forsøke å ta over EDB-arbeidet her. Foredragsholderen, som selv har vært aktiv innen språklig databehandling i en årrekke, redegjorde mot slutten av sitt innlegg for hovedtrekkene ved den tekniske utvikling av interesse for leksikografer. Han kom inn på f.eks. intelligente terminaler som kan presentere en rekke dokumenter (oppslag) parallelt på skjermen. Også utviklingen av de nye laserskrivere ble gjennomgått og tendensene innenfor optisk lesing (som ble undervurdert etter mitt syn) og fotosettingsutstyr. Videre ble utviklingen av datanett for fjern- og nærkommunikasjon omtalt, foruten hjemmedatamaskiner og søking i fulltekst databaser.

Begrensningene for leksikografene i å utnytte alt dette vil være flere, for — som Bratley sa — leksikografene er «slow, exacting, old-fashioned and poor». Ikke minst vil den svake finansielle stilling redusere farten mot det nye og stille leksikografene bak i køen i den kamp som pågår om faglært EDB-personale til slik virksomhet.

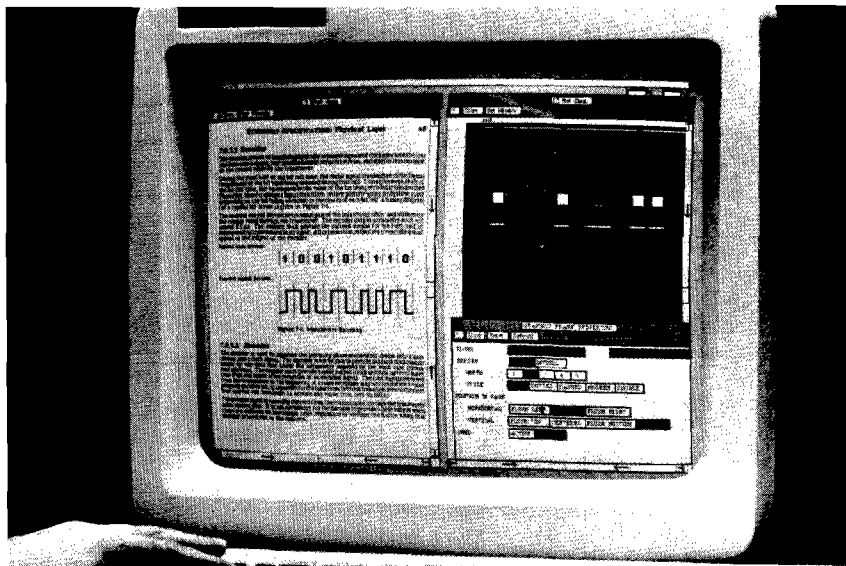
### *Presentasjonene*

På konferansen ble det holdt innpå et 30-talls foredrag og presentasjoner hvor en lang rekke europeiske ordbokstiltak ble gjennomgått. I langt de fleste tilfellene ble det orientert om hvordan leksikografisk arbeid som hadde startet ut fra konvensjonelle arbeidsmetoder, hadde forsøkt å ta i bruk EDB i løpet av prosjektperioden.

Noen titler på foredrag kan gi inntrykk av dette: *G. Gorcy*: «L'informatique et la mise en oeuvre du Trésor de la Langue Française», *N. Marinone*: «A Lexical Data Bank for Latin», *J.L. Facal*: «The Use of the Computer in the Diccionario Griego Espanol», *D.R. Howlett*: «The Use of both traditional and Computer Techniques in compiling and printing a Dictionary of medieval Latin from British Sources».

Typisk for en rekke prosjekter er at de bygger på datamaskinelt tilrettelagt tekstmateriale (evt. som resultat av andres arbeid), særlig konkordanser: *A. Cameron*: «The Dictionary of Old English, a Dictionary based on Computer Concordances» er et eksempel i så henseende.

De store vanskeligheter som vil oppstå når en skal overføre et svært leksikografisk materiale til EDB og legge om arbeidsrutinene deretter, ble presentert i *A.J. Aitkens* foredrag: «Dictionary of the Older Scottish Tongue (DOST) and Computerised Lexicography: a hopeless case?» Fra dansk side ble det orientert om hvordan en kan ta mikrodatabasemaskinteknikk i bruk i redigerings- og publiseringsarbeidet: *P.R. Petersen*: «New Words in Danish 1955-75. A Dictionary compiled and worked out in a



*Terminalarbeid med «delt skjerm» vil gi store muligheter for leksikografisk redaksjonsarbeid.*

traditional way but managed and typed via Computer».

Vesentlig forskjellig fra EDB-anvendelsene nevnt ovenfor, er de leksikografiske prosjekter der hele den leksikografiske virksomhet er basert på bruk av EDB og utformet slik at en får størst mulig utbytte av de datamaskinelle metoder.

Av denne type ble det presentert et lite — men interessant — knippe prosjekter: *Sture Allen* (Sverige): «Språkdata Lexibase System: An integrated View of a Lexical Project», *H. Zimmermann* (Tyskland): «Multifunctional Dictionaries», *N. Calzolari* og *M.L. Ceccotti* (Italia): «Description of a Database System for online Access of a Large Lexical Database» og *J.S. Bien* (Polen): «Towards Computerised Dictionaries for Morphological Languages».

Det ble også rapportert fra flere prosjekter som i sitt leksikografiske arbeid tar utgangspunkt i kommersielt tilrettelagte, maskinelle ordbøker, f.eks. som resultat av fotosetting: *J. Mullenders*: «Les Dictionnaires grammairres informatiques de l'Anglais et leur exploitation dans le projet Longman-Liege».

I et kveldsforedrag utenom programmet (kl. 22.00 — 00.30) tok prof. *Martin Kay* (Xerox Corporation) opp de fascinerende perspektiver som de nye teknologiske hjelpemiddel frembyr for leksikografisk arbeid. Kay slo fast at de nå pågående, langsiktige ordboksprosjekter (inntil 50-100 års ramme!) neppe vil bli avsluttet i det hele tatt — i alle fall ikke dersom en tenker på en omfattende ordbokspublikasjon som det naturlige slutt punktet for prosjektarbeidet.

Slik Kay så for seg fremtiden, ville ordbøker normalt ikke bli publisert i vanlig forstand. Ordbøkene ville være EDB-baserte og ved dette bli dynamiske vitenkilder i motsetning til de statiske ordsamlinger som brukes i dag. Ved bruk av EDB-metoder vil det bli mulig å anlegge helt nye perspektiver på ordstoffet og ta i bruk andre teknikker for å lette oppslag. Med EDB-baserte opplegg vil dessuten ordbøkene bli tilgjengelige for brukerne helt fra starten av.

Kay kom inn på metoder for å søke frem ord som en ikke er sikker på skrivemåten av. En tenker seg at ordet her blir skrevet inn på terminal slik en tror det skrives, og så får en det korrekte ord frem — selv i de tilfeller der det er feil stavet ved inntastingen av spørsmålet.

Med basis i eget arbeid viste Kay ved hjelp av lysbilder hvordan også mikrododateknikken i dag kan utnyttes i arbeid med et omfattende ordboksmateriale. Ikke minst vil mulighetene for å drive terminalarbeid med såkalt «delt skjerm» gi uanede muligheter for leksikografisk redaksjonsarbeid og brukerorientert behandling av opplysningene i databasen. Ved bruk av delt skjerm kan en parallelt arbeide med en rekke oppslag som er hentet fra ulike deler av f.eks. en leksikografisk datasamling.

Etter foredraget utspant det seg en livlig debatt om de farer som kunne knytte seg til en omfattende bruk av EDB-baserte vitenkilder. Enkelte deltakere uttrykte bekymring for at våre intellektuelle ferdigheter ville avta i fremtidssamfunnet. Foredragsholderen mente at det ville være nok av intellektuelle utfordringer å hanskkes med — selv om vi slapp å slå opp i ordbøker slik som vi gjør i dag.

Under konferansen ble det også gitt en interessant presentasjon av hvordan det amerikanske *National Endowment for the Humanities (NEH)* støtter ordboksarbeid. (*George F. Farr jr.*: «Supporting Lexicography at the National Endowment for the Humanities»). National Endowment for the Humanities ble opprettet i 1965 som et bevilgende organ for humanistisk forsknings- og utviklingsarbeid. For 6 år siden ble det bestemt å opprette et «*Research Materials Program*» hvor de leksikografiske prosjekter høres hjemme. Det er siden 1975 gitt støtte til over 60 ordboksprosjekter av ulik størrelse i USA. NEH har, bl.a. med tanke på den økende EDB-bruk — og for å gjøre søkerne oppmerksomme på mulighetene for databehandling — utarbeidet en serie retningslinjer som alle søkere må følge og som også regulerer forholdet til databehandling. Egne komiteer, herunder en EDB-komite, evaluerer alle prosjektene. For å sette seg grundig inn i forholdene på stedet, blir det også foretatt studiebesøk på det sted hvor prosjektet tenkes lokalisert og gitt andre former for oppsøkende konsulentbistand. Farr mente at det i USA i dag var altfor stor uvitenhet om EDB-bruk og for mye bortkastet prosjektarbeid på grunn av manglende forutsetninger for å evaluere EDB, og på grunn av unødvendig duplisering av tidligere utviklingsarbeid og systemarbeid. Tiden var nå inne, mente foredragsholderen, til å gi alle språkstudenter en elementær utdanning i databehandling. For de prosjekter som oppnår støtte fra NEH, er det gitt en rekke betingelser som dreier seg om sekundærbruk av materialet, tilgjengelighet for andre av program og EDB-opplegg og sikring av data for ettertiden. En oversikt over de ordboksprosjekter som NEH støtter, fås ved henvendelse til Senteret.

*Theodore Brunner*, leder for *Thesaurus Linguae Graecae*, University of California, Irvine, redegjorde i et innlegg for status i arbeidet for dette prosjektet, som satser på å overføre hele den klassiske greske litteratur til maskinleselig form. Til idag er ca. 42 mill. ord løpende tekst av ialt 90 mill. ord overført. Tiltaket er basert på et utstrakt samarbeid mellom klassiske filologer verden rundt — i arbeidet med å oppspore tekster og forberede dem for datamaskinell behandling.

Materialet overføres til EDB i Sør-Korea på kontrakt, og dobbelpunching blir her brukt. På konferansen ble det utdelt publikasjoner med opplysninger om tekstbankens innhold i dag. Nærmere opplysninger om den kan fåes ved Senteret.

I en etterfølgende samtale om oppbygging og drift av tekstbanker for eldre språk kretset mye av interessen rundt spørsmålet om lemmatisering av tekstene var nødvendig for de kvalifiserte brukere av denne type materiale. De kjente konflikter mellom service og en effektiv prosjektfremdrift ble også tatt opp til drøfting. Brunner var personlig meget opptatt av at det internasjonale samarbeid om distribusjon av tilrettelagt

materiale blir bedre, men så også farer ved en slik distribusjon. Siden sentrale tekster kanskje blir overført i flere ledd for ulike faglige formål, kan det oppstå ikke- autoriserte (f.eks. grafisk forenklede) varianter som bringer opphavsprosjektet i miskreditt. Andre mente det var grunn til å dempe denne uro.

### *Om lemmatisering*

En del av drøftingene på konferansen dreiet seg om behovet for å lemmatisere det tekstmaterialet man tilrettela for bruk i ordboksarbeid og ellers. Flere innlegg tok til orde for at en lemmatisering var nødvendig for å skaffe seg bedre oversikt over data og at det var viktig at en hadde klare beskrivelser for de grunnleggende metoder for den automatiske lemmatiseringen.

Det var ulike syn på hvor langt en kunne nå med automatiserte opplegg, men som en deltaker uttrykte det: «It's more easy to apply intellectual efforts afterwards», dvs. etter at maskinen har foretatt en grovlemmatisering. Generelt var det et inntrykk at deltakerne så på spørsmålet relativt pragmatisk («hva gir best resultat?») og at de ikke nødvendigvis så det som et mål å lage sofistikerte metoder for lemmatisering om like gode resultater kunne oppnås på annet vis, f.eks. ved interaktivt menneske-maskin samarbeid.

### *Om bruk av EDB i hele ordboksprosessen*

Flere deltakere var opptatt av hvordan EDB kunne tas i bruk helt fra starten av et ordboksprosjekt, og ikke minst i redaksjonsfasen. Ennå syntes den normale situasjon å være at EDB kommer inn på et relativt sent steg i mer manuelt planlagte systemer, med de begrensninger og transformasjonsproblemer som det medfører. En rekke innlegg tydet imidlertid på at man i flere europeiske land — som hos oss — etter hvert trekker EDB inn allerede ved planlegging av ordboksprosjekter. I fremtiden vil dette være det normale.

### *Vurdering av konferansen*

Som oversikten ovenfor viser, var det en stor bredde i de temaer som konferansen dekket. De store variasjonene burde det også vært tatt hensyn til organisatorisk. Ordningen med plenumsesjoner kunne opprettholdes, men hovedvekten burde vært lagt på gruppearbeid innenfor mer avgrensede temaområder. Ikke minst for dem som var interessert i bruk av datamaskinelle metoder over konkordansnivå var det frustrerende å få avsatt så lite tid til problemorientert samtale med fagfeller.

Konferansens styrke lå i at det for første gang på så bred basis var samlet leksikografer for å drøfte bruk av elektroniske hjelpemiddel i ordboksarbeid. Gjennom foredrag, innlegg og fremlagt demonstrasjonsmateriale steg konturene frem av et utrolig mangesidig leksikografisk arbeid som i dag utføres i Europa. Den aktivitet som ble dokumentert, representerer svært ulike opplegg med tanke på materiale, språk, omfang, metoder og utnyttelse av teknologiske metoder og utstyr.

Som «leksikografisk utstilling» var konferansen således vellykket. På den annen side følte trolig de aktive EDB-brukere at de hadde fått få nye impulser. En av hovedinnledderne uttrykte det slik ved oppsummeringen: «I am surprised by the lack of novelty, I have met nobody with new problems or new solutions and I have heard no whisper of the methodology of to-morrow». Og slik måtte det nødvendigvis bli: De aller fleste var aktive leksikografer som ofte — i de tilfeller EDB faktisk ble brukt — bare indirekte var i kontakt med det nye verktøy — datamaskinen.

Flere mente at en først ville få en harmonisk inkorporering av datamaskinelle metoder i leksikografisk arbeid når de leksikografiske prosjekter blir bemannet med fagpersonale med «ett bein i leksikografi og ett i datafaget», som det ble uttrykt ved slutten av konferansen.

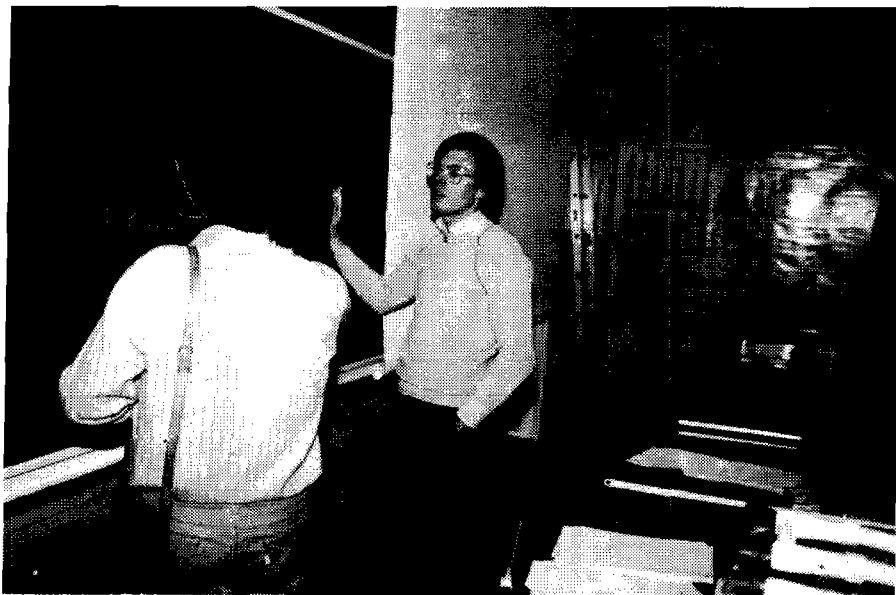
## Nordisk forskerkurs om multivariate metoder i arkeologi.

*Sigbjørn Århus*

NAVFs EDB-senter for humanistisk forskning arrangerte i perioden 17. — 27. juni et forskerkurs i bruk av multivariate metoder i arkeologi i samarbeid med statistikere og arkeologer ved Universitetet i Tromsø. Kurset samlet 18 ordinære deltakere fordelt på 5 fra Norge, 1 fra Danmark, 9 fra Sverige og 3 fra Finland. I tillegg deltok en gruppe på 4 arkeologer fra Tromsø. Programmet besto av forelesninger i arkeologi, statistikk, matematikk og EDB samt gruppeøvelser. Et terminalrom med ca 15 terminaler stod til disposisjon under hele kurset.

Arkeologiforelesningene ble holdt av dosent *Stig Welinder*, Universitetet i Oslo, professor *Ezra Zubrow*, Department of Anthropology, State University of New York, Buffalo og professor *Ian Hodder*, University of Cambridge. Dosent *Tore Schweder* og førsteamanuensis *Erik Bølviken*, Universitet i Tromsø, og forsker *Rolf Volden*, Norsk Regnesentral, sto for forelesningene i statistikk. Det ble også gitt en innføring i bruk av statistikkpakken MINITAB, som ble benyttet hyppig under kurset.

Første dag var i det vesentligste viet forelesninger i arkeologi. Hodder gav en oversikt over bruken av multivariate metoder innen britisk arkeologi, mens Welinder så på anvendelsene i Skandinavia. Tittelen på innledningsforedraget til Zubrov var «Spatial Relativism and Perception». Han kom blant annet inn på de metodiske problemene som er knyttet til grafisk fremstilling av et arkeologisk materiale. Det ble gitt eksempel på datamaskinelt fremstilte grafiske representasjoner hvor man ut fra samme



*Professor Ian Hodder, University of Cambridge, var en sentral person på det nordiske forskerkurs i arkeologi i Tromsø.*

datagrunnlag kom frem til høyst forskjellige tolkninger, avhengige av plasseringen til observatøren.

Intensjonen med kurset var å gi deltakerne en innføring i et par sentrale statistiske analysemetoder, hvor grafisk representasjon og reduksjon av data på tabell- og matriseform står sentralt. En konsentrerte seg om deskriptive og eksplorative metoder med vekt på grafisk beskrivelse, tilrettelegging og reduksjon av data.

Prinsippkomponentanalyse (PCA) og korrespondanseanalyse (CA) var hovedingrediensene i kurset. I tillegg til PCA og CA gikk man gjennom de deler av lineær algebra og statistikk som var nødvendig for å forstå de forannevnte metoder.

Mens PCA danner et naturlig utgangspunkt for matriser bestående av kontinuerlig variable, er CA velegnet til bruk på tabeller med tellingsvariable.

Det er ikke mulig her å gå inn på selve metodene. I stedet henvises interesserte lesere til et kompendium som ble utarbeidet i tilknytning til kurset. Under oppsummeringen siste dag fikk hele arrangementet en positiv omtale av deltakerne. Det ble også ytret ønske om at man om et par år arrangerer et seminar for å kunne gjøre opp ny status på feltet arkeologi og statistikk. Det er all grunn til å rose den lokale arrangementskomite i Tromsø for utmerket arbeid ved den faglige forberedelse og gjennomføring av kurset.

---

## Nordisk forskerkurs i språklig databehandling.

*Knut Hofland*

Det 5. nordiske forskerkurs i språklig databehandling ble holdt i Reykjavik fra 22/6 - 3/7 1981. Det var 20 deltakere, 7 fra Sverige, 5 fra Norge og Danmark, 2 fra Finland og 1 fra Island. Det annonserte tema for kurset var Semantiske faktorer i leksikalsk tekstanalyse. Leder var *Baldur Jónsson*. Kurset var lagt opp med forelesninger om formiddagen og øvelser på terminal om ettermiddagen. Det var to hovedforelesere på kurset, *Jerry Hobbs* fra Stanford Research Institute (SRI), California, og *Maurice Gross* fra Universitetet i Paris. Maskinen som ble brukt ved øvelsene var en VAX 11/780 og vi hadde tilgjengelig 12 terminaler.

Maurice Gross holdt forelesninger innen almen lingvistik. Han gjennomgikk syntaktiske egenskaper til enkle setninger med subjekt, predikat og objekt. Basert på en detaljert beskrivelse av egenskaper til over 7000 franske verb, ga han en oversikt over setningsstrukturregler for fransk. Verbene utgjorde ca. 50 klasser med gjennomsnitt 150 i hver klasse. For hvert verb er det beskrevet ca. 40 forhold og denne tabellen utgjør en såkalt leksikongrammatikk. Gross kom også inn på nominalisering av verb og adjektiver ved bruk av substantivering og proverb. Forholdet mellom frie setninger og idiomatiske setninger og uttrykk ble beskrevet. Eksemplene var i stor grad på fransk. Dette og manglende konkretisering (f.eks. gjennom øvingsoppgaver) gjorde at stoffet for undertegnede hang noe i luften når det gjaldt praktisk anvendelse på norsk språk.

Jerry Hobbs' forelesninger var hovedsakelig praktisk anlagt. Forelesningene var sentrert rundt SRI's DIAMOND system, et system for å definere grammatikk, leksikalsk informasjon og analyse av naturlig tekst og dialog. Systemet er opprinnelig skrevet i en spesiell LISP-dialekt, INTERLISP, men deler av systemet var overført til den variant av LISP som var tilgjengelig på VAX, nemlig Franz Lisp.

Hobbs startet med en generell gjennomgang av LISP og gikk så over til funksjonene for definisjon av leksikon og grammatikk. Hovedtrekkene i parsingalgoritmen (bottom up) ble skissert. Allerede 2. dag var deltakerne i stand til å legge inn egne grammatikker. Han gikk videre og beskrev hvorledes det kunne legges restriksjoner på reglene slik at det ble samsvar mellom leksikalske og semantiske egenskaper.

Tilslutt kom Hobbs inn på overgang fra analyse-tre til logisk form og hvorledes en ved predikatlogikk kunne trekke slutninger basert på tekster og en kunnskapsbase. Dette siste fikk en ikke prøvet ut i praksis.

I tillegg til disse forelesningene holdt *Sture Allén* to forelesninger. Den første gjaldt semantiske aspekter i leksikalsk tekstanalyse. Han gjennomgikk på hvilke nivåer semantiske kriterier ble brukt under arbeidet med Nusvensk frekvensordbok. Prosjektet Lexikalisk databas ble også presentert. Den andre forelesningen beskrev området språklig databehandling ved hjelp av den inndeling som var gjort ved innkalling til konferansene COLING og ALLC, samt definisjoner som sentrale personer innen feltet har gitt. Videre gjennomgikk han kravene til et «Language User Program System».

*Bente Maegaard* og *Hanne Ruus* beskrev arbeidet med EF-kommisjonens oversettelsessystem. Oversettelsen er brutt ned i tre trinn: analyse, overførsel og generering. Analysen gir et syntakstre med merkelapper som kommer fra analysen og fra ordboken. Disse merkelappene er sentrale ved generering.

*Greges Koch* orienterte om hvorledes han hadde brukt program for logisk analyse til behandling av naturlig språk.

Kurset fikk en noe mer dreining mot syntaksanalyse enn det som var satt opp som kurstittel, i alle fall i de praktiske øvelsene. Men det var en stor fordel at deltakerne hadde et ferdig system fra starten som en kunne bruke ved øvingene. Baldur Jónsson og hans assistenter stod for et utmerket arrangement.



*Første gang EDB er drøftet på Tingvellir?*

---



# NORDDATA 81

*Eirik Lien*

Denne årlige begivenheten var i år lagt til København med Danmarks Tekniske Højskole i Lyngby som arrangementsted. NordDATA er den nordiske sammenkomsten av alle aktive datafolk, uansett bransje. Det fører selvsagt til at det blir svært stor deltakelse (i år i overkant av 2000), og foredrag som spenner over hele EDB-spekteret. Lista over hovedtemaer vil belyse det: Ledelse og administrasjon, Systemanalyse, Programmering, EDB-drift, Databaseteknikk, Datakommunikasjon, Utdannelse, Anvendelse, Kontorautomasjon og EDB og samfunn.



I alt ble det presentert ca. 200 foredrag på to og en halv dag. Å gi en uttømmende oversikt over det faglige innholdet vil være umulig. I stedet vil jeg prøve å gi mitt inntrykk av den plassen humanistene hadde i dette bildet og kommentere de foredragene som grenser inn på vårt område.

To foredrag av de 200 hadde tema knyttet til EDB-anvendelser i humanistiske fag: *Bente Maegaards* om automatisk orddeling og mitt om bruk av databaseteknikk for å systematisere sammengjengse bibeltekster. Det er vel kanskje symptomatisk for forståelsen av den plassen humanistene har i EDB-bildet at vår sesjon var plassert under hovedtema «Kontorautomasjon», riktignok med undertittel «Sprogbehandling». Det er vel også symptomatisk at de to foredragene samlet henholdsvis 35 og 20 tilhørere av de 2000.

En del andre foredrag kom inn på våre interesseområder, bl.a. to under temaet «Relasjonsdatabaser»: «On semantical problems when using query languages» og «To nye metoder til humansproglige databaseforespørsler». Begge disse tok opp problemer knyttet til bruk av naturlig språk i kommunikasjon med datamaskin.

Under temaet fotosetting, som er interessant for oss som tekstprodusenter, ble det presentert to foredrag. Det ene behandlet en maskinuavhengig måte å beskrive det typografiske oppsettet av en tekst på, kalt TEXUS, det andre var en gjennomgåelse av et dokumentframstillingssystem (PHOTODOC) som brukes ved RECKU, Københavns Universitet

Hovedinntrykket fra NordDATA bekrefter at vi er fåtallige i selskap med realister, teknologer, økonomer og administratorer. Vi har kanskje ikke den store appellen i et kommersielt miljø. Men jeg tror det er viktig at vi markerer oss i bildet, at vi viser at datamaskinen kan brukes til andre formål enn lønnsutbetaling, tekniske beregninger og varedistribusjon.

Benta Maegaards foredrag viste således hvordan lingvisten kan gi bidrag til å løse problemer i tekstbehandlingssystemer. Derfor er det viktig at også humanistene deltar på slike konferanser og bidrar med foredrag, spesielt når vi kan få anledning til å komme i tale med EDB-folk utafør vår krets. Vi bør vise at vi naturlig hører med i det store og varierte EDB-miljøet.

Humanistene har også en god del å hente på slike arrangementer. Mange av foredragene er av allmenn interesse, både de som tar opp nye bruksområder og teknikker og de som behandler den virkningen EDB har på samfunnet. Og det er nyttig å få vite hva andre EDB-folk er opptatt av.

# The 6th. International Congress of Applied Linguistics.

## *Per Vestbøstad*

The International Congress of Applied Linguistics gjekk av stabelen 9.-14. august i Lund, Sverige. Kring 1000 deltakarar fra 64 land — dei fleste språklærarar eller forskarar - kunne velja mellom godt over 300 presentasjonar i form av plenumsforedrag, spesialforedrag, arbeidsgrupper og korte prosjektpresentasjonar.

Mellom plenumsforedraga nemner eg *Roger Shuy*, University of Georgetown, Washington D.C.: «Can linguistic evidence build a defense in a criminal law case?» og *Aaron Cicourel*, University of California, San Diego: «Language and the structure of belief in medical communication».

Spesialforedraga tok m.a. opp emne som «Computer aids to translation», «Quantitative contrastive analysis» og «The teaching of spoken language».

Emne i arbeidsgruppene var t.d. «Language projects of the Council of Europe» og «The language of Public Documents».

Dei om lag 290 korte presentasjonane var fordelte slik på ulike emnegrupper:

1. First language learning and teaching (18)
2. Methods in second language teaching (18)
3. Progression in second language learning (24)
4. Error analysis and contrastive linguistics (28)
5. Evaluation and testing (20)
6. Teaching materials, text books and pedagogical grammars (12)
7. Language teaching and technology (8)
8. Language for spesific purposes (14)
9. Translation and interpreting (21)
10. Terminology and lexicography (16)
11. Production and comprehension (33)
12. Communication problems in society (28)
13. Bilingualism and immigrants' language problems (26)
14. Language planning and society (25)

Ettersom alle 14 emnegruppene vart presenterte samstundes, kunne deltakarane naturleg nok berre få med seg ein brøkdel av alt dette, og det var ofte vanskeleg å velja mellom interessante emne. Men når det i god tid før kongressen var sendt ut «preprints» til alle korte presentasjonar, hadde vi eit godt grunnlag å velja på, samstundes som vi også hadde eit konsentrat av dei presentasjonane vi ikkje nådde over.

Datamaskinell språkhandtering hadde ingen stor plass ved konferansen, men under seksjonane for *læring* og *teknologi* og *terminologi* og *leksikografi* fanst der likevel ein del presentasjonar innom dette området.

Eg vil særleg nemna franskmannen *Demaiziere* (Université Paris VII) som rapporterte frå eit prosjekt der ein hadde utvikla eigne forfatterspråk til kurs som nyttar datamaskinstøtta læring. Dermed har ein betre reidskap til å analysera verbale svar, slik at dialogen elev — maskin vert naturlegare.

Kongressen gav eit overveldande inntrykk av kor vidtfemnande fagområdet «anvendt lingvistikk» er. Underskrivne merkte seg særskilt at mange forskarar er opptekne av talemålsanalyse og -innlæring, og av meningsformidling.

## Computer Corpora in Research and Teaching

NAVFs EDB-senter for humanistisk forskning arrangerte 1.-3. juni et internasjonalt seminar om EDB og språkforskning. Seminaret, som var lagt til Bryggens Museum, Bergen, samlet 20 deltakere fra Sverige, England, Nederland, India og Norge.

Å forske i språk kan i dag være svært forskjellig fra det som var vanlig bare for noen få år tilbake. Mens den tradisjonelle språkforskeren brukte mye tid til å samle inn sitt materiale, kan man i dag ved datamaskinens hjelp hurtig få fram et omfattende forskningsmateriale og dermed bruke mer tid til den rent språklige analysen. Forutsetningen er imidlertid at det fins maskinlesbare tekster og effektive metoder for automatisk søking og gjenfinning.

NAVFs EDB-senter for humanistisk forskning har arbeidet aktivt med slike spørsmål, både nasjonalt og internasjonalt. Senteret er bl.a. sekretariat og operativt organ for ICAME (International Computer Archive of Modern English). Denne organisasjonen samler og distribuerer informasjon om engelsk språkmateriale som er tilgjengelig for elektronisk databehandling og informerer om språkforskning bygd på det innsamlede materiale.

På seminaret i Bergen ble følgende hovedtemaer tatt opp:

- prinsipper for innsamling av tekster
- metoder for språklig analyse
- metoder for automatisk søkning i teksten
- arbeidets betydning for språkundervisningen på universitetsnivå

Selv om den enkelte forsker kan ha god nytte av maskinlesbare tekster, kreves det mye tid og arbeide for å tilrettelegge tekstene og å utvikle metoder for effektiv utnyttelse av dem. Internasjonalt samarbeid er i den forbindelse helt avgjørende.

Ved seminaret ble det presentert flere sentrale prosjekter som er av stor interesse både for engelsk språkforskning og for språkforskning generelt i Norge og andre land.

Det er planen å samle de fleste foredragene i bokform.

*Interesserte kan henvende seg til*

*NAVFs EDB-senter for humanistisk forskning,  
postboks 53, 5014 Bergen-Universitetet.*



*Noen av deltakerne på seminaret om EDB og språkforskning på Bryggens Museum 1.—3. juni 1981.*

## Tre glade hospitanter.

I tida 21. til 30. april hadde Senteret besøk av tre stipendiater som gjennomgikk et opplæringsprogram i bruk av EDB til drifts- og forskningsoppgaver. Tilbudet var beregnet på driftsmedarbeidere innen den kontor-tekniske stillingsgruppen. Internt betegnes opplegget som *hospitantprogrammet*, og det skal etter planen bli et årlig tilbud.

Målsettingen med kurset er å gi deltakerne et oversyn over de ulike steg i EDB-arbeidet, øving i terminalbruk og opplæring i enklere former for elektronisk databehandling på relevant faglig materiale. Det var bl.a. tilrettelagt øvingsdata for to av de tre institusjonene som var representert.

Hver arbeidsdag var delt i to: før lunsj var det innledninger og orienteringer fra Senterets EDB-konsulenter, tida etter matpause ble brukt til øvingsarbeid på terminal og til kontakt med fagmiljøer i Bergen. Deltakerne fikk også en orientering i bruk av mikrodatamaskin, bl.a. med vekt på tekstbehandling.

Humanistiske Data har slått av en prat med de tre og forhørt seg litt om deres bakgrunn og yrke og hvilke erfaringer de tar med seg fra kurset.

*Jane Floor* er registreringsassistent ved Arkeologisk museum i Stavanger. Hun registrerer forminndata på EDB i et fellesprosjekt mellom Arkeologisk museum, Rogalandsforskning og Rogaland interkommunale datasenter. Stillingen medfører bl.a. ansvar for publisering av registreringsrapporter. Hun har relativt lang erfaring med bruk av EDB.

*Anita Paulsen* er kontorfullmektig ved Nordmøre museum i Kristiansund, og har som en av sine oppgaver å registrere museumsdata på katalogkort for optisk lesning. Mange av opplysningene gjelder vår fiskerihistorie og arkeologiske funn fra den såkalte Fosna-kulturen. Museet samarbeider med Senteret i form av betalte oppdrag. Høyt på ønskelisten til museet står egen terminal-tilknytning eller egen mikrodatamaskin.

Anita Paulsen har ikke hatt noen form for EDB-opplæring før dette kurset, men følger et innføringskurs pr. brevscole på heimstedet.

*Else Marie Vedå* har i over 5 år arbeidet med tilrettelegging av data for prosjektet «De kulturgeografiske registreringer på Vestlandet». Via terminal til Univac 1100 ved Universitetet i Bergen registrerer og kontrollerer hun data om forskjellige kulturgjenstander: redskap, klær, kister o.s.v. Til nå er det foretatt over 30.000 slike registreringer. Hun har tidligere gjennomgått et innføringskurs i EDB.

Hva fikk man så ut av kurset? Alle tre understreker at — til tross for den relativt store spredning i forkunnskaper ble utbyttet meget godt. Det hele var så fleksibelt og individuelt lagt opp at forskjellen i erfaringer og kunnskaper ikke spilte noen stor rolle. De understreker imidlertid at Senteret bør sammensette mer homogene grupper i fremtiden, og dette har vi tatt til etterretning. Særlig ved litt større hospitantgrupper vil dette være av betydning.

Paulsen satte særlig pris på den betydning praktiske øvinger hadde hatt. Det gjorde at problemene kom inn både gjennom hodet og fingrene. For Vedå hadde kurset ført til at mange ting hun «durte på» i det daglige arbeidet, nå fikk sin forklaring. Floor understrekte bl.a. betydningen av den kontakt med andre fagmiljøer i Bergen som Senteret hadde formidlet.

Ellers hindrer vår beskjedenhet oss å nevne at alle tre roste Senterets personale for vennlighet og tålmodighet med den enkelte.

Det kan til slutt nevnes at *oppfølging* er et problem både i forbindelse med dette



*Fra venstre: Else Marie Veddå, Anita Paulsen og Jane Floor.*

hospitantprogrammet og med stipendiatopplegget for vitenskapelig personale. De tre hadde en slags edb-messig «ut-i-kulden»-følelse når de nå skulle hjem til sine respektive arbeidsplasser. Det var derfor et sterkt ønske om opplegg for egenstudier, videre kontakt med Senteret osv. Vi beklaget at dette for tiden ikke var mulig å tilby innenfor Senterets arbeidsprogram, men at en skulle se spesielt på problemet i forbindelse med neste års program. Bl.a. ble det nevnt at en litteraturliste for videre lesning kunne utarbeides.

## Utviklingsseminar

En av de sentrale arbeidsoppgaver for EDB-konsulentene i Senteret og ved universitetene er å veilede andre. På møter og konferanser, i brev og notater, ved telefonsamtaler og besøk lærer en fra seg det en måtte kunne om de forskjellige EDB-emner. Men en forutsetning for å undervise andre er selvsagt at en har lært noe sjøl og hele tiden lærer noe. Dette var også bakgrunnen for det såkalte *utviklingsseminar* på Ustaoset i slutten av mars i år. Foruten fagpersonalet i Senteret deltok de fire EDB-konsulentene ved universitetene og 7-8 andre som står sentralt i det humanistiske EDB-miljøet. Til sammen var det 18 personer som i tre dager benket seg bak hotellgardiner som effektivt skjulte påskesol og raske løyper utenfor.

Et sentralt tema på seminaret var programmeringsspråket PASCAL. Innlederen *Jan Olav Hauge* fra Universitetet i Tromsø oppga tre grunner til at en burde lære seg PASCAL: 1) det er «i skuddet», dvs. språket blir stadig mer populært i undervisning og utvikling av programsystemer, 2) språket er svært anvendelig og alminnelig på mikrodatamaskiner, 3) språket er godt egnet til strukturert programmering, dvs. er raskere å skrive, mer oversiktlig og lettere å flytte fra en maskin til en annen enn mange av de tidligere språk. Hauge viste så noen egenskaper PASCAL har v.h.a. et program innenfor språklig databehandling. I en egen innledning om *strukturert programmering* viste han ved eksempler og sammenlikninger hvorfor PASCAL er et egnet språk til den type programmering. Kombinasjonen strukturert programmering/ PASCAL kan bli særlig verdifull for humanister.

Mikrodatamaskiner var det andre hovedtemaet på seminaret, og *Børre Stenseth* var invitert som foredragsholder. Han er amanuensis ved Østfold distriktshøgskole der de har valgt å konsentrere sin undervisning og forskning omkring denne type maskiner. Han holdt et engasjert «pro-foredrag» som ble etterfulgt av en like engasjert «pro-et-contra»debutt. Stenseth pekte på de mange fordeler ved mikromaskiner, bl.a. at den tar liten plass, er rimelig og har et stort bruksområde. Som sentrale bruksområder ble nevnt tekstbehandling, arkivfunksjoner, statistikk/regnskap, undervisning, adressering m.v. Av de tre bruksmulighetene: en-bruker, fler-bruker og nettverk var Stenseth tilhenger av det siste, dvs. et nettverk av en-brukermaskiner der hver person har ansvar for og kontroll med sine egne data.

*Sigbjørn Århus* spurte i sin innledning om mikrodatamaskiner var spesielt aktuelt for humanistisk forskningsarbeid, og tok utgangspunkt i erfaringene med Senterets egen mikromaskin.

I en egen innledning tok Stenseth for seg programmering og organisering av data på mikrodatamaskinen.

Spørsmål og debatt dreide seg i stor grad om forholdet mellom store maskiner og mikromaskiner. De fleste var enige om at mikromaskinen er et ypperlig redskap til sitt bruk, og har et stort bruksområde, men at en må ha både store og små maskiner, avhengig av hvilke arbeidsoppgaver som skal utføres.

Andre temaer på seminaret var tekstbehandling, kvalitetsutskrift, programpakke for språkstatistikk og arbeidet med en håndbok i EDB for humanister.

*Knut Hofland* innledet om tekstbehandlingssystemer kombinert med generell databehandling. Han formulerte ca. 20 krav til et godt tekstbehandlingssystem, og vurderte mikromaskiner og store maskiner ut fra disse kravene. *Eirik Lien*, *Ivar Fønnes* og *Sigbjørn Århus* tok for seg noe av det utstyr som i dag finnes for kvalitetsutskrift: matriseskriverer, skriverer med faste typer og laserskriverer.

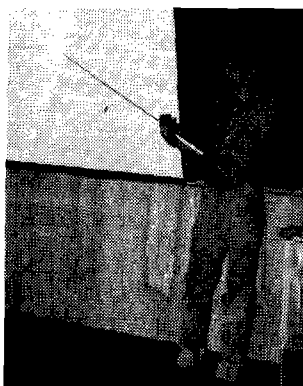
*Roald Skarsten* drøftet behovet for og eventuelle krav til en god programpakke for språkstatistikk. Det var enighet om behovet for et slikt tilbud, og det ble skissert en arbeids- og tidsplan for det videre arbeid.

Til slutt presenterte Eirik Lien en del tanker om et kompendium i EDB for humanister. Også her ble det skissert en handlingsplan for det videre arbeid.

Det foreligger referat fra alle innledninger og debatter. Disse er samlet i en rapport som er gratis og kan bestilles ved henvendelse til Senteret.



*Jan Olav Hauge fra Universitet i Tromsø holdt foredrag om PACAL og strukturert programmering*



*Børre Stenseth fra Østfold distriktshøgskole holdt foredrag om mikromaskiner.*

# MELDINGER

## Forskere knyttet til Senteret

Som det er redegjort for i Humanistiske data nr. 1-81 (side 69) har Rådet for humanistisk forskning opprettet to forskerstipend knyttet til Senteret.

Fra 1.9.81 er dosent *Stig Welinder*, Oldsaksamlingen, Universitetet i Oslo engasjert for å arbeide med edb og kvantitative metoder i arkeologisk forskning. Welinder vil sammen med Senteret i første rekke arbeide med en programpakke for arkeologisk forskning med vekt på statistiske analyser. I tillegg vil han arbeide med et eget forskningsprosjekt hvor edb-metoder blir benyttet. Welinder vil i samarbeid med Senteret dessuten gjennomføre en forskningsorientert seminarvirksomhet, både på nasjonal basis og rettet mot en enkelt faginstusjon.

*Per-Kristian Halvorsen*, Ph.D, vil fra 1.1.82 arbeide i Senteret med oppgaver innenfor språklig databehandling. Halvorsen, som for tiden arbeider ved Massachusetts Institute of Technology (MIT), Center for Cognitive Science, vil sammen med Senteret gjennomføre et forskningsprosjekt innen feltet automatisk syntaktisk og semantisk analyse.

Prosjektet vil foregå som ledd i arbeidet mot et brukerorientert system for anvendelse av naturlig språk ved kommunikasjon med EDB-systemer (f.eks. informasjonssystemer). I tilknytning til forskningsarbeidet vil det bli arrangert spesialseminar innenfor området.

## EDB og filosofi.

2. og 3. november arrangerer Senteret et seminar om bruk av EDB innen filosofi. Siktemålet med seminaret er i første rekke å belyse hvordan elektronisk databehandling kan tas i bruk ved sammenstilling og analyse av filosofiske tekster - i dette tilfellet Wittgensteins forfatterskap. Som innledning til dette tema vil det bli gitt en orientering om bruk av EDB ved behandling av vitenskapelig kildemateriale. Også erfaringer fra EDB-bruk i et forskningsprosjekt om utviklingen av Bergens-skolen i meteorologi vil bli presentert (Ralph Jewell). Initiativtakerne til seminaret er professorene Viggo Rossvær og Knut Erik Tranøy.

Seminaret holdes i Bergen og vi kommer tilbake med rapport i neste nummer av bladet.

## De nordiske datalingvistikkdager Trondheim 22-24 oktober 1981

Som ved de to tidligere arrangementene i Gøteborg i 1977 og København i 1979 er formålet å utveksle ideer og erfaringer i bruk av datamaskinelle metoder for å løse lingvistiske problemer. Lingvistikk er her ment i vid betydning.

Utvekslingen skjer gjennom foredrag med tilknyttet diskusjon og gjennom demonstrasjoner av programverktøy, men også gjennom uformelle samtaler. Det blir satt av tid både til det formelle og uformelle.

Videre informasjon blir sendt ved henvendelse til: *Eirik Lien, EDB-tjenesten for humanistiske fag, Universitetet i Trondheim, N-7055 DRAGVOLL, NORGE. Påmelding til samme adresse.*

---



## **COLING 82 - call for papers**

The Ninth Conference on Computational Linguistics. Date: July 5th-10th, 1982. Location: Prague, Czechoslovakia Sponsored by: International Committee on Computational Linguistics in association with: Linguistic Institute of L. Stur, Slovak Academy of Science, Bratislava and Faculty of Mathematics and Physics, Charles University, Prague.

Papers are invited for presentation especially from the following domains:

- theories, methods and problems of computational linguistics
- relations of computational linguistics to computer science, mathematics, linguistics, artificial intelligence etc.
- representation of knowledge and inferencing as they relate to language understanding
- applications of natural language processing, information, retrieval, communication, machine translation and machine-aided translation, automatic understanding of texts, speech recognition and synthesis, etc.

*Authors wishing to present a paper should submit 4 copies of a 3 to 4 page summary, double spaced, by December 1st, 1981, to COLING 82, MFF UK, Linguistics, Mainstranske n. 25, 118 00 PRAGUE 1, CZECHOSLOVAKIA*



## **EDB-konferanse for humanistiske fag ved de regionale høyskoler, 5. — 6. januar 1982.**

NAVFs EDB-senter for humanistisk forskning inviterer representanter for de humanistiske fagmiljøer innen det regionale høyskolesystem til en konferanse om bruk av edb i humanistisk forskning. Henvendelsen går i første rekke til alle distriktshøyskoler og til de pedagogiske høyskoler som tilbyr års- eller halvårseininger innen humanistiske fagområder.

Konferansen har som mål å samle fagmedarbeidere i humanistiske fag til en drøfting av edb-spørsmål med tilknytning til deres fagområder.

*Konferansen blir holdt i Bergen 5-6 januar 1982*

Den første dagen vil bli viet mer generelle edb-emner:

- Sentrale trekk ved databehandlingsarbeid
- Edb-utstyr av særlig interesse for humanister
- Databehandling av vitenskapelig kildemateriale
- Oversyn over feltet humanistisk databehandling
- Programbehov i de humanistiske fag
- Organer for konsulenthjelp og utviklingsarbeid innen humanistisk databehandling.

Den andre dagen tenkes arrangert som en konkret drøfting av edb-bruk innenfor språkfag, litteratur, historie osv. Det vil bli lagt vekt på presentasjon av pågående og utført prosjektarbeid, data og programmer. Deltakerne vil bli inndelt i faggrupper. Det vil bli lagt vekt på å drøfte hvordan fagmedarbeidere ved de regionale høyskoler kan utnytte mulighetene for databehandling i sitt arbeid.

Deltakerne forutsettes å dekke reise- og oppholdsutgifter. I særlige tilfeller kan arrangøren kunne dekke kostnadene. Vi er derfor også interessert i påmelding fra deltakere som må ha ekstern finansiering av reise og opphold.

Nærmere opplysninger om konferansen gis av:

**NAVFs EDB-senter for humanistisk forskning**

**Harald Hårfagresgt. 31**

**Boks 53**

**5014 BERGEN-Universitetet**

**telefon (05)21 00 40**

**Søknadsfrist 15. november.**

## Leksikografi og datateknologi

Det synes som bruk av datamaskinelle metoder i leksikografisk arbeid vies stadig større oppmerksomhet (jfr. referatet fra Pisaseminaret annetsteds i dette nummer). EF-kommisjonen er et av de organer som i dag satser sterkt på bruk av EDB i ordboksarbeid og i virksomhet knyttet til EDB-støttet oversettelse og terminologisk arbeid.

Den 7. — 9. juli ble det holdt et internasjonalt symposium i Luxembourg med tittelen «Lexicography in the Electronic Age». Symposiet inneholdt sesjoner om «Creation of lexicons», «New technologies», «Term banks», «Publishing and the future use of lexicons». *Humanistiske Data mottar gjerne referat for publisering fra eventuelle norske eller nordiske deltakere.*

---

## Symposium on Artificial and Human Intelligence.

A Symposium on Artificial and Human Intelligence, sponsored by the Human Factors Subcommittee of NATO, will be held at the Chateau de Chapeau Cornu near Lyon, France, October 26-30, 1981. The two major topic areas to be addressed are: (1) the structure of memory, the representation of the problem, and the knowledge and belief available to the problem-solver — key issues in both human and machine intelligence, and (2) the structure of language and the role that languages play in problem solving. Other areas of interest include inductive, deductive, and analog thinking, roles in research of expert systems, benchmarks and yardstick problems, constraints of complexity and computability, and priorities in AI research.

*Further information may be obtained from the conference director: Dr. Alick Elithorn, The Royal Free Hospital, Bond Street, London NW3 2QG, ENGLAND*

---

## ALLC International Meeting - call for papers

The ALLC International Meeting and ACM for 1981 will be held at Oxford University on 17-18 December. Papers will be given on the afternoon of Thursday, 17th December and all day on Friday, 18th December. The meeting will finish with a sherry reception in St. Cross College followed by a conference dinner in the hall of St. John's College on the evening of Friday 18th December.

Attendance at the meeting is open for non-members. The themes of the conference will be

- (1) Textual editing techniques;
- (2) Methods of publication of computer-assisted library and linguistic research.

Residential accommodation has been reserved in St. John's College, Oxford, from lunch on 17th December until breakfast on 19th December inclusive, but the College has asked us to ensure that all enquiries should be addressed to *ALLC International Meeting, Oxford University Computing Service, 13 Banbury Road, Oxford OX2, 6 NN, UK.*

The cost of attendance at the meeting is expected to be in the region of 50 pounds, including accommodation and meals, sherry party and conference dinner.

## **EDB-kurs for museumstilsette. 30. november - 11. desember 1981.**

I samarbeid med *Norske Kunst- og Kulturhistoriske Museer (NKKM)* vil NAVFs EDB-senter for humanistisk forskning i Bergen skipa til eit EDB-kurs for museumspersonale i tida 30. november - 11. desember 1981.

Kurset rettar seg i første rekkje mot museumspersonale som er ansvarleg for registrering og katalogisering av gjenstandsmateriale. Museum som skal ta i bruk NKKMs EDB-kort eller som på annan måte i ferd med å nytta EDB i katalogiseringsrutinar, vil ha førerett til kurset. Talet på deltakarar er avgrensa til 12.

På kurset vert det førelesingar med almenn innføring i bruk av EDB. Deltakarane vil dessutan få ei innføring i bruk av NKKMs EDB-kort og orientering om praktisk røynsle med bruk av EDB i museumssamheng. Det vil verta lagt stor vekt på praktiske øvingsoppgåver. Ein vil nytta anlegget ved Universitetet i Bergen, Univac 1100/82. Det vil dessutan verta høve til å bruka mikrodatamaskin (ALTOS) og å nytta universitetsanlegga i Oslo, Trondheim og Tromsø via UNINETT. Deltakarane må sjølve dekkja utgifter til reise og opphald. I spesielle tilfelle kan NAVFs EDB-senter gje stønad til reise- og opphaldsutgifter. Senteret kan ordna med innkvartering på hotell (kr. 2.250 for full pensjon, kr.2000 for rom m/frukost). Dei som ønskjer å ordna slik innkvartering, må gjera merksam på det i søknaden.

Nærare opplysningar om kurset kan ein få ved å venda seg til NAVFs EDB-senter, førstekonsulent *Sigbjørn Århus* tlf. (05) 210040/2956 eller til konservator *Magne Velure*, Hardanger Folkemuseum, 5797 Utne, tlf. (054) 66927.

*Søknad vert å senda til:*

*Norske Kunst- og Kulturhistoriske Museer.*

*Pilestredet 15*

*OSLO 1*

---

*Conference on Data Bases in the Humanities and Social Sciencies blir holdt i Pisa i Juni 1982.*

*The 6th International Conference on Computers in the Humanities blir arrangert ved North Carolina State University i 1983.*

---

## **EDB-tjenestene ved universitetene i 80-årene.**

Universitetsrådets komite for EDB-samarbeid startet i mai 1979 en utredning om de funksjoner EDB-tjenestene på universitetene burde tillegges i 1980-årene. Arbeidet ble fullført i januar 1981 og rapporten er nå til høring i institusjonene.

Hensikten med utredningen var å

- gi en oversikt over vesentlige trekk i den teknologiske utviklingen
- peke på hvordan nyere teknologi kan nyttes til å skape bedre tjenestetilbud
- skissere en modell for den videre utbygging av EDB-tjenestene innen og mellom universitetene
- skissere konsekvenser av en slik modell for brukerne, EDB-sentrene, budsjetteringen og samarbeidet mellom universitetene
- foreslå tiltak for å bedre tjenestetilbudet i det neste ti-året

EDB-tjenestene ved universitetene har siden starten vært organisert rundt EDB-sentra som fungerer som felles, operative EDB-miljøer for hele universitetsområdet. Det har i den siste tiden vært reist spørsmål om EDB-sentrene som organisatoriske enheter er nødvendig lenger i sin nåværende form.

I rapporten kommer det klart frem at universitetenes EDB-systemer i 1980-årene vil fordeles på mange og ulike typer ressurser. Forbedret ytelse og fallende priser på maskinvaren vil føre til stadig økende bruk av selvstendige, større og mindre datamaskiner i undervisning og forskning. Egne, kraftige maskiner for den enkelte arbeidsplass vil etterhvert bli tatt i bruk. Samtidig vil mange ressurser være felles, for hele institusjonen eller for mindre enheter og grupper. Dette fordi enkelte ressurser fremdeles vil være så kostbare at de krever felles økonomiske løft, eller vil bli rettet mot felles oppgaver, datagrunnlag eller programvare. *Å få dette spektrum av mange ulike typer ressurser til å spille sammen som en fullverdig tjeneste for alle universitetets brukere av databehandling, vil være EDB-sentrenes utfordring i årene fremover.*

Grunnlaget for et samvirke om felles tjenester vil være bruk av kommunikasjon mellom ressursene for overføring og utveksling av data gjennom kommunikasjonsnettene (datanett). Omfattende prosjekter er i gang nasjonalt (UNINETT) og internasjonalt (ISO) for å definere standardene. Slike sammensatte systemer av ulike typer ressurser er i rapporten gitt navnet «*åpne systemer*»: de er åpne for alle del-systemer som følger standardene. Modellene for universitetenes EDB-tjenester i årene fremover vil være den åpne system-modellen, og utvikling, vedlikehold og drift av kommunikasjonsnettene og felles-komponentene i denne vil bli en viktig del av EDB-sentrenes oppgaver.

Fallende pris på maskinvarer i forhold til ytelsen vil føre til at programvaren, med vedlikehold og videreutvikling, overtar en økende del av kostnadene i et EDB-system. Teknologien for bedre programutvikling står nærmest på stedet hvil i forhold til maskinutviklingen. Samtidig vil større deler av den mer anvenderorienterte del av programarbeidet foregå i brukermiljøene. Utstrakt bruk av ferdige programpakker og spesialpråk orientert mot anvendelse vil gjøre dette mulig. Å anskaffe, videreutvikle og vedlikeholde denne programvaren mot maskinvare, og å kunne flytte programmene til de til enhver tid best egnete ressurser, enten de er felles eller enkelt-stående, vil kreve EDB-tjenester som for et universitet med fordel bør være en fellestjeneste. Sammen vil disse felles, generelle programvaretjenester være EDB-sentrenes hovedoppgaver fremover.

I rapporten gis det en rekke interessante (og ikke-tekniske) presentasjoner av viktige sider ved den pågående og fremtidige teknologiske utvikling. Rapporten anbefales derfor for alle som vil sette seg inn i bruk av EDB på universitetene i årene som kommer.

*Tankevekkende er f. eks. opplysningene om at mens programutrustningskostnadene i 1965 lå på ca. 10% av totalkostnadene ved utvikling av et nytt system, ventes de tilsvarende kostnader i 1985 å bli ca. 80%!*

For humanistiske forskere vil det økende spekter av datatjenester innenfor dataregistrering og datapresentasjon gjøre databehandlingen stadig mer formålstjenlig. Det tenkes her f.eks. på de datagrafiske teknikker og utviklingen av brukervennlige kommunikasjonsformer med datamaskin i mest mulig naturlig språk. Også introduksjon av utskriftsutstyr som gir tilnærmet boktrykkskvalitet og høy grad av fleksibilitet bør nevnes. Tekstbehandling forventes å bli standard hjelpemiddel både innen administrasjon og forskningsarbeid.

*Rapporten kan skaffes ved henvendelse til universitetenes EDB-sentra.*

# BIBNETT

BIBNETT er et samarbeidsprosjekt mellom BIBSYS (bibliotekautomatiseringsprosjekt), Norsk dokumentdata, Universitetsbibliotekene i Oslo og Bergen, UNINETT og Riksbibliotekjenesten. Prosjektets størrelse er ca. 3 årsverk pr. år.

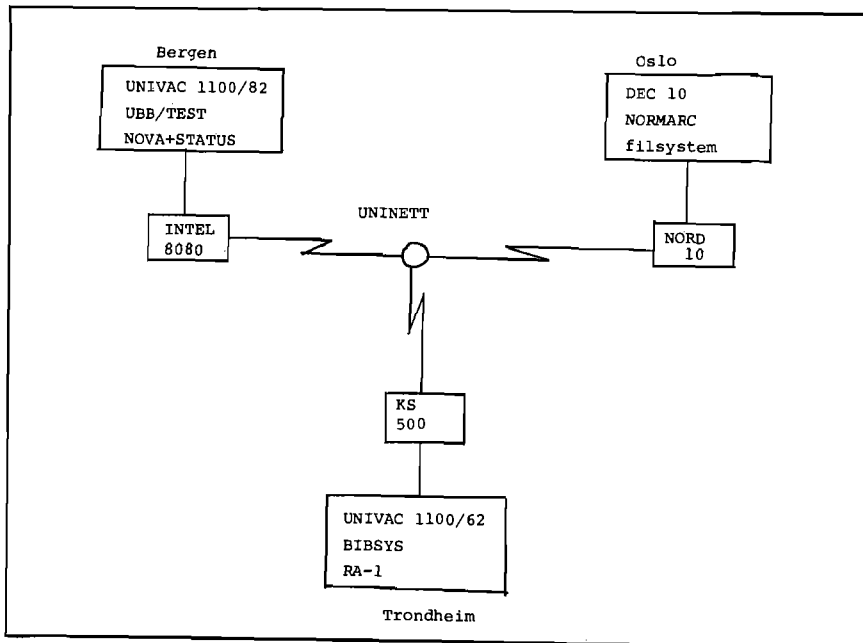
Formålet med BIBNETT er å utprøve og påvirke UNINETT's egenskaper for biblioteksformål og å avklare om program-til-program kommunikasjon mellom uavhengig utviklede programsystemer på forskjellige datamaskiner er teknisk og økonomisk realistisk i Norge.

Tre biblioteksystem skal kommunisere med hverandre ved hjelp av UNINETT. Biblioteksystemene BIBSYS, NORMARC og UBB/TEST kjøres på forskjellige maskiner UNIVAC og DEC, og benytter forskjellige databasesystemer, RA-1, DEC's filsystem og NOVA\*STATUS..

Det skal søkes etter og overføres bibliografiske poster. Systemene presenterer seg for en annen datamaskin som en terminal og sender innloggings- og oppstartingsdata over nettet. Det er utarbeidet en applikasjonsprotokoll for biblioteks- og informasjonssystemer.

Vi har en program-til-program kommunikasjon i prøvedrift. Deler av applikasjonsprotokollen er implementert, og videre implementering vil fortsette i hele prøvedriftsfasen frem til 15/11-81. Deretter vil vi vurdere de resultater vi er kommet frem til. Prosjektet BIBNETT avsluttes 31/12-81.

*Åse Støgard*



## Senterrapport nr. 19.

### Statistiske metoder på arkeologisk materiale.

Rapporten bygger på en konferanse i november 1980 arrangert i Bryggens Museum av NAVFs EDB-senter for humanistisk forskning. På seminaret deltok 18 spesielt inviterte areologer, EDB-konsulenter og fagstatistikere, og hensikten var bl.a. å få et sjekkpunkt for hvor vi står på dette feltet i dag. Konferansen var også en del av forberedelsene til det nordiske forskerkurs om statistiske metoder i arkeologi som seinere ble arrangert i Tromsø (se referat s. 28).

Rapporten består av sammendrag av 13 innledninger foruten åpningsforedrag, oppsummering, program, deltakerliste og et planleggingsnotat om det nordiske forskerkurset i Tromsø. Rapporten kan bestilles ved henvendelse til Senteret. Pris kr. 35. ISBN 82-7283-017-5 (mars 1981).

## Senterrapport nr. 20.

### EDB-prosjekter i de humanistiske fag.

Som det ble redegjort for i Humanistiske Data 1-80, har Senteret kartlagt EDB-bruken i de humanistiske fag gjennom en spørreskjemaundersøkelse ved årsskiftet 1979-80.

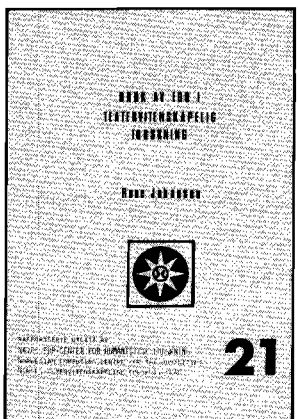
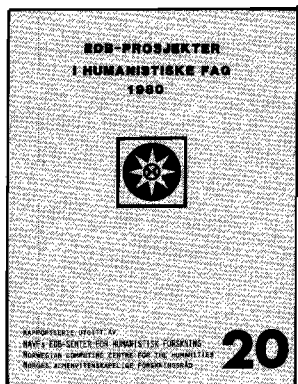
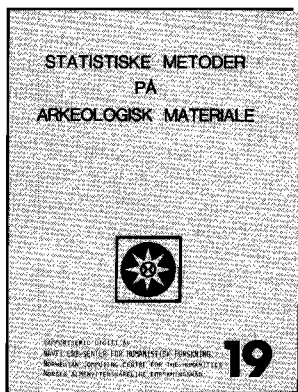
Gjennom direkte henvendelse til prosjektlederne høsten 1980 har vi sikret oss deres samtykke til å offentliggjøre resultatene i en egen prosjektrapport og samtidig fått opplysningene om prosjektene ajourført.

Samtidig vil vi benytte anledningen til å nevne at Rådet for humanistisk forskning i samarbeid med NAVFs Utredningsinstitutt og Senteret vil utgi en prosjektkatalog i løpet av høsten som dekker all humanistisk forskning i Norge i dag. Prosjektkatalogen er en ajourføring av Humanistisk forskning 1977. Den vil på flere punkter kunne utdype det bilde av EDB-virksomheten i de humanistiske fag som Senterets prosjektoversikt gir. *Rapport nr. 20 kan bestilles ved henvendelse til Senteret. Pris kr. 45. ISBN 82-7283-018-3 (juni 1981). De som har levert bidrag får rapporten gratis tilsendt.*

## Senterrapport nr. 21.

### Bruk av EDB i teatervitenskapelig forskning.

Rapporten er skrevet av NAVF-stipendiat *Rune*



Johansen og bygger på den bruk av EDB han har gjort i sitt prosjekt «Norsk scenedekorasjonskunst 1908 - 1935. Tradisjon og «modernisme» belyst gjennom arbeider av Jens Wang og Oliver Neerland».

Rapporten skisserer først det faglige innhold i prosjektet, og beskriver deretter hvordan EDB er brukt for å ordne materialet. Til slutt drøftes hvordan teaterarkivalier generelt kan ordnes ved hjelp av EDB, til beste både for teaterinstitusjonene og forskningsmiljøene. Rapporten kan bestilles ved henvendelse til Senteret. Pris kr. 35. ISBN 82-7283-019-1 (mai 1981)

## Senterrapport nr. 22.

### Årsmelding 1980.

Senterets årsmelding forelå i mai og er distribuert til ca. 700 aktuelle adresser. Den er på 24 sider og inneholder opplysninger om prosjektsamarbeid, kurs- og konferanser, informasjonstiltak, program- og metodeutvikling foruten generelle opplysninger om Senteret og virksomheten i 1980. Det er laget et fyldig engelsk sammendrag. Årsmelding er gratis og kan bestilles ved henvendelse til Senteret.

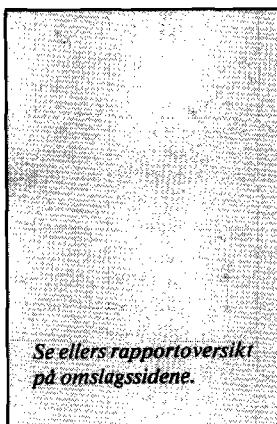
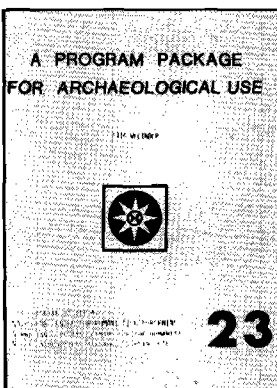
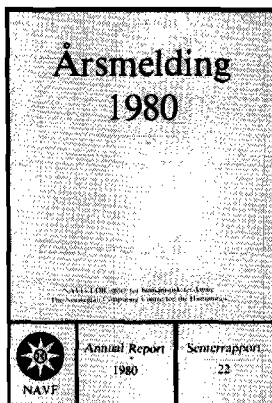
## Senterrapport nr. 23.

### A program package for archaeological use. Stig Welinder.

«The present package of computer programs is aimed at the statistical and numerical processing of archaeological data. It consists of programs based on the SPSS package, (Nie et al. 1975; referred to as «SPSS» in the continuation), one program borrowed from the Stockholm University computing centre, and programs written especially for the package by the author in collaboration with *Ivar Fonnes*, computer adviser at the Faculty of Arts at the University of Oslo..

The package contains programs for the following methods: Descriptive statistics including bivariate analysis, Principal components analysis, Factor analysis, Cluster analysis, 2 methods of seriation.

The ambition of the programs is that an analysis of an archaeological source material should end in a graphic representation of the data and their inherent structure.» Rapporten kan bestilles ved henvendelse til Senteret. ISBN 82-7283-021-3 (Oslo, 1981). Pris kr. 45,—.



# SUMMARY

## **FONETIKK OG STATISTIKK.**

### **Phonetics and Statistics.**

In this article Prof. Knut Fintoft, Department of Linguistics, University of Trondheim, describes the close interrelation between Phonetics and Statistics in modern research work.

As an academic discipline Phonetics intersects with other specialised fields, e.g. Psychology, Sociology, Pedagogy and Medicine. But irrespective of orientation, data collection in one form or another is crucial to the analysis of phonetic phenomena. The meaningful analysis of such data involves well adapted methods for quantification. It is here that computational techniques come to play an important role in the field.

Prof. Fintoft mentions that for these reasons all students studying Phonetics at the University of Trondheim are introduced to data processing techniques.

By a series of examples Prof. Fintoft shows how quantitative methods are used in the study of the frequency of speech sounds and in the identification of the frequency of vowel formants.

Moreover we are shown how statistical analysis may be needed to perform perception tests, and to produce synthetic speech sounds and also to determine the relationship between long and short vowels.

In the field of Sociolinguistics statistical methods are similarly basic resources for, among other things, keeping under control the considerable number of variables brought into discussion by this field of study.

## **DATABEHANDLING OG HUMANIORA.**

### **Computers and the Humanities.**

The article by the Director of the Centre, Jostein H. Hauge, gives a general preview of a series of interviews and presentations which will appear in the ensuing issues of *Humanistiske Data*.

The aim of this series is to provide a background for examining the need within the humanities for computer-related education and to present an overview of the different forms of education and training offered by the Norwegian universities today.

The author contends that Norway is generally lagging behind many other industrialized nations with regard to general education in data processing. While such education in many countries already has been, or is rapidly being integrated in the curricula of the ordinary school systems, the Norwegian school authorities are only on the point of recognizing the need for general knowledge of the new technologies. Students entering the Arts Faculties thus normally have no previous knowledge of computational methods.

The reason for introducing data processing courses in the humanities are seen as threefold:

1. Knowledge of computers and communication technology will be needed in the future to understand some of the basic mechanisms of society. A fair amount of knowledge hence will be a prerequisite also for a competent participation in the general debate on societal and cultural issues engendered by the ubiquity of computers and computer-related technologies.
2. Knowledge of computers will broaden the professional platform of teachers in the humanities. Acquainted with computational methods the better would they be equipped to promote a broad view of the societal and ethical consequences of the new technologies.



3. In the future knowledge of computers will be necessary for those students who aim at a career in research. Computers are ostensibly more and more becoming «naturalized» research tools in the humanities. The author also claims that computer-oriented researchers have special responsibilities towards the extra-mural community. One example is that of Linguists who have an obligation to contribute to the development of communication forms with computers based on natural language and generally to bridge the «linguistic» gap between the technological hardware and its users.

But there is yet another — a more matter of fact — reason why researchers in the humanities should pay more attention to computational techniques, namely the need to comply with government efforts to make the use of public resources (e.g. research) more cost-effective. In the future the latter desideratum may well turn out to be a powerful agent of change in humanistic research — even in an oil nation like Norway.

## **INTERVJU MED IVAR FONNES.**

### **Interview**

In this interview Ivar Fønnes, Faculty ADP-consultant to the Faculty of Arts, University of Oslo, describes the newly established one-year course in Computing in the Humanities. The course is a response to the growing interest among Arts students for computational techniques. In History, for example, a number of students make extensive use of computers in the preparation of their final theses (hovedfagsoppgaver). Up to this time students and researchers from the humanities have been forced to take courses given by other faculties, where they are often alienated by the use of exercises and examples fetched far from their fields of knowledge and lines of interest. Basic humanistic data problems and analytic strategies will be at the very core of the new course.

## **RAPPORTER**

### **Reports**

A series of reports is given of international and national conferences and courses.

*Jostein H. Hauge* reviews the workshop on the Possibilities and Limits of the Computer in producing and publishing Dictionaries, in Pisa, May 20 — 22. The workshop — held under the auspices of the Standing Committee of the Humanities, European Science Foundation (ESF), was attended by approximately 50 participants from 15 European countries and the U.S.A.

The answers to a questionnaire distributed by the organizers at a preparatory stage in an effort to find out the state of the art, show that well over 50 per cent of the current lexicographic enterprises in the field in some way or other actually make use of data processing. The conference revealed, however, that relatively few of the projects so far have integrated the new techniques from the initial stages of the lexicographical work.

*Eirik Lien* reports from the largest annual computer conference in the Nordic countries — NORDDATA. In Copenhagen last June some 2000 participants were given a choice of more than 200 papers — among them only 2 relating to computer projects in the humanities (one of them given by the author). Mr. Lien stresses that humanists ought to make themselves better heard also at such general computer conferences, both in the roles of participants and speakers.

*Knut Hofland* gives a report from the Nordic research course in computational linguistics held in Reykjavik June 23 — July 3. This time the course was oriented towards automatic syntactic and semantic analysis of natural language. The principal guest lecturers were Prof. Maurice Gross, University of Paris, and Prof. Jerry Hobbs, Stanford Research Institute. The 20 participants (from all the five Nordic countries) were given ample time to practise the use of the computational systems presented in the lectures.

Reports are also presented from the National Computer Conference (NCC) in Chicago in May 4 — 7., and the joint conferences, the Fifth International Conference on Computers in the Humanities, and Easier and More Productive Use of Computing, Ann Arbor, May 17 — 22. Participants from the Centre gave three papers at the Ann Arbor Conferences.

In addition to these international events there has been a series of seminars and courses conducted by the Centre. In June an international seminar on Computer Corpora in Research and Teaching was held in Bergen with participants from England, India, the Netherlands, Sweden and Norway. The seminar dealt with present international computer work on English Text Corpora with a view to facilitating the exchange of experience. Different ways of exploiting text materials in Computational Linguistics and Linguistics proper were also taken up by the Bergen seminar. A selection of contributions will later appear in book form.

A Nordic research course in the uses of multivariate methods in Archaeology was held in Tromsø June 16 — 27. The course was organized by the Centre in collaboration with archaeologists and statisticians at the University of Tromsø. Principal lecturers were Prof. Ian Hodder, University of Cambridge, and Prof. Ezra Zubrow, New York State University. 20 archaeologists from 4 Nordic countries took part in the course where a considerable time was spent on correspondence analysis and on data from archaeological field work which was attacked by the help of the statistical techniques presented at the course.

## MELDINGER

### News

In the course of the Autumn the Centre will arrange a seminar on the applications of computers in practical and theoretical theatre work, with emphasis on information retrieval, another seminar on text analysis of philosophical texts and a course on methods for establishing computational archives in cultural museums.

A recent plan for the future development and co-operation between the computing centres of the four universities in Norway is described, and finally there is a brief mention of current use of the computer networking system operating between the universities (UNINETT) for on-line bibliographic services (BIBNETT).

Vennligst send meg:

\_\_\_\_\_ eks. av Rapport nr. \_\_\_\_\_ á kr. \_\_\_\_\_

\_\_\_\_\_ eks. av Årsmelding 1980 (gratis)

\_\_\_\_\_ eks. av Humanistiske Data 1-81 (gratis)

\_\_\_\_\_ eks. ICAME NEWS nr. \_\_\_\_\_ (gratis)

Jeg ønsker å abonnere på Humanistiske Data

Jeg ønsker å abonnere på ICAME NEWS.

Navn \_\_\_\_\_

Institusjon \_\_\_\_\_

Adresse \_\_\_\_\_

*forts. fra 2. omslagsside:*

Rapport nr. 13. *Datatjenester for og datasamarbeid mellom kunst- og kulturhistoriske museer.* Februar 1980. ISBN 82-7283-010-8. Pris kr. 30,

RAPPORT nr. 14. *NOVA\*STATUS, systemdokumentasjon. Brukerveiledning.* 2. opptrykk februar 1980. ISBN 82-7283-011-6. Pris kr. 15,

RAPPORT nr. 15. *Ivar Fønnes: Tekstsøking på tegnnivå.* Januar 1980. ISBN 82-7283-012-4. Pris kr. 15,-.

RAPPORT nr. 16. *Årsmelding 1979, NAVFs EDB-senter for humanistisk forskning.*

RAPPORT nr. 17. *Svein Lie: Automatisk syntaktisk analyse. Del 1. Grammatikken.* Desember 1980. ISBN 82-7283-014-0. Pris kr. 30,-.

RAPPORT nr. 18. *Datateknologi og humanistisk forskning. Bidrag til en NAVF-utredning.* Desember 1980. ISBN 82-7283-015-9. Pris kr. 30,-.

RAPPORT nr. 19. *Statistiske metoder på arkeologisk materiale* ISBN 82-7283-017-5. *Mars 1981. Pris kr. 35.*

RAPPORT nr. 20. *EDB-prosjekter i humanistiske fag 1980.* ISBN 82-7283-018-3 Juni 1981. Pris kr. 45.

RAPPORT nr. 21. *Bruk av EDB i teatervitenskapelig forskning.* Rune Johansen. ISBN 82-7283-019-1. Mai 1981. Pris kr. 35.

RAPPORT nr. 22. *Årsmelding 1980. NAVFs EDB-senter for humanistisk forskning.* Gratis.

RAPPORT nr. 23. *A program package for archaeological use.* Stig Welinder. ISBN 82-7283-021-3. August 1981. Pris kr. 45.

---

for humanistisk forskning  
Harald Hårfagesgt. 31, Boks 53,  
5014 Bergen.

Professor dr. jur.  
**MONS SANDNES NYGÅRD**  
UNIVERSITETET I BERGEN  
Altegt. 34, 5014 Bergen, Universitetet  
Telefon: (05) 21 2940  
Privat: Postboks 24, 5068 Hjeltnes  
Telefon: (05) 226124

BERGEN den 25. April 1981

Bergmanns Boktrykkeri A/S,  
Bergen.

Underskrivne har nyleg hatt til trykking ein  
var utarbeidd hos Bergmanns Boktrykkeri A/S. Manuskriptet  
til trykkeriet ved datamaskinell teknikk, og vart overført  
telesamband, og nytta til datastykt sats. Etter ommodning  
vil eg gjerne gje uttrykk for at denne framgangsmåten er  
eit fullt tilfredsstillande resultat.

Trykkinga galdt ein vitenskapleg kjeldebublikasjon  
av domprotokollar frå slutten av 1600-talet, Overhoffrotts-  
domar 1 (1667-79). Hovudinnhaldet er originale skriftstykke  
frå dette tidsrom med vanskeleg ortografi, både bokstav- og  
ordkombinasjonar. Vanleg satteteknikk ville ha blitt svært  
fattande korrekurarbeid. Med den valde framgangsmåten kunne  
dette for ein stor del snarast inn. Ein viktig føremøn var  
likevel at det heile tida fram til trykkeklar sats var høve  
til å halde kontroll med overføringa og utarbeidinga av  
trykksatsen, takk vere eit nært samarbeid med trykkeriet.  
Den oppsette tidsplanen for trykkinga vart punkteleg fylgt.

Med helsing

*Mons Sandnes Nygård*  
Mons Sandnes Nygård

Til: Fotosats  
Fra: Datamaskin

**bergmann**

Telefon med modem 31 27 42

Telefon 31 65 20

Øvregaten 17, 5000 Bergen