

Jann Scheuer: Hans hustru og hendes bryster (uredigerede uddata)

Den Danske Ordbog

Undersøgelse af forbindelser med *hans* og *hendes*

	HANS	Klasse	T-score	HENDES	Klasse	T-score
1	kone	s	30.16	mand	s	-28.86
2	hustru	s	12.72	bryster	s	-10.08
3	folk	s	11.72	verden	s	-9.57
4	disciple	s	8.87	hår	s	-9.27
5	sønner	s	8.25	øjne	s	-8.86
6	brødre	s	7.61	mor	s	-8.05
7	værker	s	7.29	kinder	s	-7.55
8	mænd	s	7.29	krop	s	-7.36
9	opfattelse	s	6.81	stemme	s	-7.21
10	politiske	adj	6.77	kjole	s	-7.00
11	pik	s	6.77	forældre	s	-6.92
12	død	s	6.55	far	s	-6.91
13	kolleger	s	6.53	majestæt	s	-6.67
14	venner	s	6.13	ben	s	-6.61
15	billeder	s	6.06	kusse	s	-6.06
16	eller	kon	6.05	ansigt	s	-5.96
17	parti	s	6.04	mands	s	-5.89
18	regering	s	5.99	elsker	s	-5.72
19	musik	s	5.99	hud	s	-5.47
20	tilhængere	s	5.62	skød	s	-5.38
21	politik	s	5.46	veninder	s	-5.16
22	egen	adj	5.36	lille	adj	-5.07
23	film	s	5.31	lår	s	-5.06
24	kammerat	s	5.28	trusser	s	-4.55
25	indsats	s	5.28	ægtemand	s	-4.44
26	kontor	s	5.20	smil	s	-4.42
27	ideer	s	5.18	mund	s	-4.39
28	trofasthed	s	5.17	veninde	s	-4.30
29	firma	s	5.11	taske	s	-4.26
30	eget	adj	5.07	læge	s	-4.22
31	rige	s	4.99	bluse	s	-4.14
32	tid	s	4.97	børn	s	-4.12
33	værk	s	4.93	varme	adj	-4.01
34	produktion	s	4.91	søstre	s	-4.00
35	afgang	s	4.78	dør	s	-3.91
36	land	s	4.75	hånd	s	-3.90
37	medarbejderes		4.66	ældste	adj	-3.78
38	ord	s	4.58	barn	s	-3.76
39	styre	s	4.57	hvide	adj	-3.43
40	hold	s	4.50	hals	s	-3.43
41	efterfølger	s	4.48	latter	s	-3.42

Mette Kunøe og Erik Vive Larsen (udg.):
5. Møde om Udforskningen af Dansk Sprog,
Århus 1995

UDSKRIFTER AF DANSKE SAMTALER

Præsentation af nyt korpus med udskrifter af danske samtaler

Af Jakob Steensig (Aarhus Universitet)

1. Introduktion

På Institut for Lingvistik i Aarhus er vi nogle studerende og ansatte som de sidste to-tre år har arbejdet med analyser af danske samtaler. Vi er nu ved at oprette et egentligt data-korpus, som skal indeholde bånd og udskrifter af de samtaler vi efterhånden har fået optaget. Korpus kommer til at hedde *LAUS* (*L*ingvistik *A*arhus *U*niversitet, *S*amtalekorpus), og det er meningen at det skal stå til rådighed for forskere og studerende i hele landet.

Formålet er at udvide datagrundlaget for empiriske studier af dansk talesprog. Vi forestiller os at korpus kan bruges fx af forskere som har udviklet hypoteser på basis af egne data, og som gerne vil afprøve disse på yderligere data, af studerende som mangler samtaledata til opgaver eller projekter, af lærere som gerne vil bruge autentiske eksempler, m.fl.

2. Praktiske oplysninger om *LAUS*

Korpus kommer til at indeholde et indeks og en række bånd med tilhørende udskrifter. Til hver optagelse vil der foreligge et *datablad* med en kort beskrivelse af deltagere, optagelsesomstændigheder m.m. Generelt vil man kunne få adgang til både bånd og udskrifter mod at garantere anonym behandling af data og mod et løfte om at stille evt. nye udskrifter til rådighed for *LAUS*. I enkelte tilfælde kan der være knyttet specielle anonymiseringskrav til bestemte optagelser; det vil i så fald fremgå af databladene.

Vi har søgt penge til medhjælpstimer til at udarbejde datablade og evt. et søgeprogram, give databladene et fælles format, samle udskrifter og evt. tilpasse dem vores minimale udskriftsnormer (se nedenfor), og til at komplettere de udskrifter der allerede er lavet. Hvis vi får pengene vil korpus blive tilgængeligt i løbet af foråret 1995. Nærmere oplysninger kan fås ved henvendelse til undertegnede på Institut for Lingvistik, Aarhus Universitet, 8000 Århus C, tlf.

8942 2176, fax 8619 1699, e-mail: linjs@stud.hum.aau.dk.

3. Beskrivelse af data

De data som bliver tilgængelige er i første omgang optagelser fra 11 samtaler, som alle foregår på dansk mellem modersmålstalende voksne danskere. Situationerne er dagligdags uformelle samtaler samt enkelte private møder. Ved indsamlingen er der blevet lagt vægt på at samtalerne skulle være så lidt påvirkede af observationen som muligt, og der er derfor i næsten alle tilfælde tale om samtaler som ikke er arrangerede til ære for optagelsen, men hvor vi har fået lov at optage en situation som ville have fundet sted alligevel. I de tilfælde hvor data er arrangerede i et eller andet omfang, fremgår dette af databladene.

3.1. Datablade

Når man søger i korpus efter data, vil den nemmeste indgang være at bruge databladene. På Fig. 1 nedenfor kan man se et eksempel på hvordan et datablad kan se ud:

Samtalebetegnelse: Kvote 2
Optagelsesdato: Marts 1994
Optager: Eva Husted Nielsen
Anvendelsesbetingelser: Anonymisering af personnavne
Længde og placering i samtalen: 45 min., uden start og slutning
Deltagere:
- antal: 3
- køn: 1K + 2M
- alder: i 20'erne
- forhåndsbekendtskab: læser samme fag på samme niveau, kender hinanden godt.
- uddannelses- og arbejdsmæssig baggrund: studerende
Setting:
- medium: face-to-face
- sted: Århus, i den ene M's kollektiv, i stuen.
- genstande: teudstyr
- deltagerne placering: ved et bord. En deltager er af og til i køkkenet
- deltagerne formål: uformelt møde om bedømmelse af ansøgere på deltagerne uddannelsesinstitution.
Optagelsessituationen
- autencitet/oplæg: ikke eliciteret. Deltagerne har kendskab til mikrofonen og formålet med optagelsen.
- tekniske data: Philips N2233 (alm. kassetteoptager), ekstern kondensatormikrofon med kugleprofil
- mikrofonplacering: på bordet
- optagerens placering: deltager i samtalen
- andet: Deltagerne er med af interesse for emnet.
Udskrifter:
- længde: 3 min. 15 sek.

- placering på båndet: 3.27-3.42; 4.50-5.20
- anvendt til: Emneopgave i samtaleanalyse på Inst. for Lingv. AU
- referencer: Nielsen, Eva Husted 1994: *Så går vi rundt om en uenighedsbusk!* - Gør vi? Århus: Inst. for Lingv., AU.
Andre bemærkninger:

Fig. 1: Eksempel på datablad.

Der er ikke tale om et endeligt format. Vi mangler bl.a. at diskutere mulighederne for at udarbejde et søgeprogram til databladene, og at tage stilling til hvilke krav et sådant i givet fald vil stille til de kategorier som bruges på bladet.

Men eksemplet i Fig. 1 giver et billede af de oplysninger vi ønsker at have med. Til bladet kan flg. bemærkes:

Vi foreslår at man bruger vores samtalebetegnelse + "LAUS" ved reference til samtaler i korpus. I det viste eksempel hedder data altså "Kvote 2, LAUS".

Som det ses, har vi relativt få etnografiske/sociale oplysninger i forhold til hvad man kræver i fx "ethnography of speaking" (jf. fx Saville-Troike 1982) eller sociolingvistiske undersøgelser (jf. fx Gregersen & Pedersen (eds.) 1991). Dette skyldes at data ikke er indsamlet med etnografiske eller sociolingvistiske formål for øje. Men nogle af de oplysninger vi har om deltagerne uddannelsesmæssige eller sociale baggrund, om talebegivenheden, om deltagerne forhold til hinanden eller anden diskurshistorie, er angivet på databladet.

Databladet giver også oplysninger om hvad data er blevet anvendt til. Det vil være muligt at få kopier af de referencer der opgives, og vi forventer så til gengæld at evt. nye artikler eller opgaver bliver føjet ind på databladet.

3.2. Udskrifter

Udskrifterne bliver tilgængelige på papir og på diskette. De er foretaget på Macintosh-computer i programmet *syncWriter*, som er specielt udformet til udskrifter af samtaler i *partitur*, dvs. hvor hver taler har én linie hele vejen igennem (jf. Fig. 2 nedenfor):

[DK1, LAUS]

1 Carsten	er det de der halvfjerdserbrune låger, (.)
Emil	
Pia	mng -t (h) jeg ka ikke

- 2 Carsten
Emil hehehe[he.]
Pia huske hva farve de hæ:r, men de: 'ikke pæne. [ehe] .hng
- 3 Carsten
Emil [-hh]
Pia [så:] der sådan et et ø::h krydsfinerskab til kopper, (.)
de:t
- 4 Carsten
Emil
Pia simpelthen det 'første der ska ske det er at det ska ud,

Fig 2: Eksempel på partiturudskrift

Vi udarbejder et sæt minimumskonventioner som udskrifterne kan tilpasses, men denne tilpasning er ikke en forudsætning for at data kan være med i korpus. Minimumskonventionerne kommer til i et eller andet omfang at lægge sig op ad konversationsanalysens konventioner, sådan som det ses i eksemplet i Fig. 2 (se iøvrigt Atkinson & Heritage 1984 for en introduktion til disse). I minimumskonventionerne vil en relativt stor nøjagtighed blive tilstræbt hvad angår angivelse af pauser, tøven og udtale af ikke-fonemiske lyde. I nogle udskrifter vil der derudover være grundigere angivelser af fx intonation, lydstyrke, stemmekvalitet m.m. Med hver udskrift følger en beskrivelse af de anvendte konventioner.

Som det ses af ovenstående har vi valgt ikke at lægge os op ad konventionerne i *Dansk Standard* (Gregersen 1992). Denne beslutning er taget efter at vi har afprøvet både dette og andre udskriftssystemer, og har fundet at konversationsanalysens konventioner (eller noget der ligner dem) passer bedst til det vi laver.

4. Hvad kan LAUS bruges til?

Vi forestiller os som nævnt at studerende og lærere kan gå til LAUS og finde samtaler og eksempler til opgaver og undervisning. Men vi mener også at et korpus som LAUS kan få forskningsmæssig betydning. Et samtalekorpus gør det muligt at efterprøve resultater og hypoteser fra egne analyser på et større materiale. Ofte vil et interessant fænomen kun optræde i ringe antal i en enkelt samtale, og en forståelse af fænomenet kræver at det er muligt at gennemgå og analysere flere tilfælde af det pågældende fænomen. De resultater som den amerikanske konversationsanalyse (se fx Atkinson & Heritage 1984, Heritage 1985), og den svenske Initiativ-respons analyse (Linell & Gustavsson 1987) har

opnået, havde ikke været mulige uden gennemgange af sådanne større samlinger af samtaledata. Det er bl.a. denne mulighed vi nu ønsker at tilvejebringe for danske samtalers vedkommende.

På Mødet om Udforskning af Dansk Sprog var der mange interessante oplæg om samtaleanalyse, og mange af dem formulerede problemstillinger som med fordel kunne efterprøves i et samtalekorpus. Jeg nævner her et par stykker, som faldt mig ind under nogle af de spændende fremlæggelser på mødet:

Erik Møller gav i sit oplæg "Den prototypiske mundtlige fortælling" en meget udførlig definition af "den konversationelle fortælling". Den er udarbejdet på basis af det enorme korpus som findes i Bysociolingvistikprojektet ved Institut for Dansk Dialektforskning, Københavns Universitet. Data er dér sociolingvistiske interviews og arrangerede gruppesamtaler med deltagelse af observatør. Et korpus som LAUS ville gøre det muligt at se om Erik Møllers definition også dækker fortællinger i ikke-arrangerede dagligdags samtaler.

Mie Femø Nielsen talte om "vi-etablering ved afdelingsmøder" ud fra sit spændende korpus af audio- og videooptagede afdelingsmøder i større virksomheder. Hun kom bl.a. ind på hvordan deltagere eksplicit ændrer deres formuleringer så de passer med den normerede sprogbrug i afdelingen. Hun gav eksempler på formuleringer som fx (citeret efter hukommelsen): "Vi har ansvaret for- vi er forhandlingsansvarlige for.." og "tilpasning - det er de ord de bruger" og viste hvordan deltagerne ved deres brug af sådanne formuleringer forholder sig til afdelingen som referencepunkt og lederen som den der har hovedansvaret for at definere, legitimere og fortolke. Et tjek i et korpus som LAUS ville kunne vise om lignende formuleringer også forekommer i dagligdags samtaler uden udnævnte ledere, og kunne evt. give anledning til en præcisering af hvorfor samtaledeltagere i det hele taget ekspliciterer deres valg af begreber.

Marianne Johansen forelagde nogle resultater fra sit spændende arbejde "Om samtaleemners organisering". Hun har fundet frem til en række "emnehandlingskategorier" som er baseret på en analyse af både semantiske propositioner og interaktionen. En af dem er den "reproducerende" emnehandling, som iflg. Marianne Johansen har en "præ-lukkende" funktion. Dette er en ny og dristig hypotese i forhold til det klassiske arbejde omkring lukninger af sekvenser (fx

Schegloff & Sacks 1973). Men Johansens definition af emnehandlinger synes så præcist operationaliseret sprogligt og interaktionelt, at det må være muligt at gå ind i andre data (som fx LAUS) og undersøge om reproducerende emnehandlinger også dér optræder (efterfulgt af "kollaborerende" emnehandlinger) i forbindelse med lukning af emner.

Dette blot et par hurtige eksempler som blev inspireret af mødet. Jeg vil nedenfor kort fortælle om vores egen baggrund for at lave LAUS og om hvad materialet hidtil er blevet brugt til.

5. Baggrund

Baggrunden for initiativet til LAUS er først og fremmest gode erfaringer med at samarbejde om data. Dette samarbejde er foregået dels på instituttet i forbindelse med et kursus om samtaleanalyse i foråret 1994, dels blandt instituttets medarbejdere som har brugt hinandens data som udgangspunkt for diskussioner, og dels i "Arbejdskreds Fremmedsproglig Interaktion", som er et samarbejde mellem (fortrinsvis jyske og fynske) forskere der analyserer fremmedsprogede samtaledata, og hvor arbejdet er baseret på at man analyserer deltagernes data på skift.

Data i LAUS stammer fra ovennævnte kursus og fra undertegnede s arbejde (jf. Steensig 1993, 1994a, 1994b). De er blevet brugt til opgaver om uenighed i samtaler (Nielsen 1994), om præference-forhold i ytringspar (Pedersen 1994), om reparation (Tolstrup 1994), og om turtagningssignaler (Ladefoged 1994, Ladefoged & Ilcus 1993, Steensig ovennævnte). Disse opgaver tager alle fat på områder som ikke før er blevet indgående behandlet på danske data. De kan fås ved henvendelse til instituttet.

6. Nogle kommentarer til diskussionen ved fremlæggelsen

Jeg havde ikke troet at fremlæggelsen af dette initiativ ville give anledning til nogen særlig debat. Men det viste sig at der var temmelig mange indvendinger på mødet mod måden vi greb det an på. Nedenfor vil jeg gerne tage fat på et par af disse indvendinger. Jeg vil understrege at min fortolkning af indvendingerne og mine kommentarer til dem står for egen regning.

Nogle mødedeltagere undrede sig over at vi spillede ud med sådan et initiativ uden først at kontakte Institut for Dansk Dialektforskning, som iflg. en beslutning fra 1986 af forskere i hele landet er udnævnt til centrum for talesprogsforskningen i Danmark. Dette vidste jeg ikke. Og det er egentlig underligt, da jeg har deltaget i mange ph.d.-seminarer og konferencer hvor samtaleanalyse har været diskuteret, uden at nogen har gjort mig opmærksom på at der fandtes et sådant centrum.

Der blev også udtrykt forundring over at jeg forkastede udskriftskonventionerne i *Dansk Standard*. Dette er til gengæld gjort med vilje, idet jeg som sagt har afprøvet disse konventioner og ikke været tilfreds med dem. Pladsen her tillader ikke en nærmere diskussion af dette, men i næste nummer af Aarhus Universitets sprogvidenskabelige tidsskrift, *SPRÅU*, vil jeg gøre rede for mine overvejelser vedrørende udskriftskonventioner for samtaleanalyse, og jeg håber også at dette problem kan blive gjort til genstand for egentlige diskussioner på fremtidige konferencer om emnet. Men jeg oplever ikke at forskellige udskriftskonventioner har været et problem i de situationer hvor jeg har deltaget i diskussioner af andres data, og jeg har også gode erfaringer med at oversætte andres udskrifter (i fx *Dansk Standard*) til mine egne udskriftskonventioner.

I det hele taget tror jeg ikke på *ét centrum* for samtaleanalyse eller på *ét fælles sæt konventioner* for udskrifter. Samtaleanalysen i Danmark må nødvendigvis være *polycentrisk*. Der bliver lavet samtaleforskning mange forskellige steder i landet (på vistnok alle universiteter, på Lærerhøjskolen, på mindst én teknisk skole osv.). Vi arbejder inden for forskellige metodologiske paradigmer, med forskellige datatyper og med forskellige samarbejdspartnere i udlandet. Det specielle ved situationen i Danmark er at vi har en livlig diskussion på tværs af paradigmer, og det synes jeg er positivt og noget vi skal udvikle. En af tilgangene til sådanne diskussioner kan være at vi ser på og bruger af hinandens data.

Vi ønsker ikke med LAUS at udgøre nogen form for centrum. Der er snarere tale om at vi lægger vores data "ud på nettet", fordi de nu engang findes og relativt let kan gøres tilgængelige. Snarere end at se LAUS som et forsøg på at lave noget centralt, vil jeg håbe at det kan ses som en opfordring til andre om at gøre noget lignende. Jeg er nemlig overbevist om at der ligger guldgruber af data rundt omkring, som vil kunne frigives med relativt små anstrengelser.

Jeg kom ved fremlæggelsen til at bruge en betegnelse, "den jysk-fynske akse", som faldt nogle mødedeltagere lidt for brystet. Jeg vil derfor benytte lejligheden til at understrege at jeg med denne - lidt spøgefulde - betegnelse blot mente det ovenfor nævnte samarbejde i "Arbejdskreds Fremmedsproglig Interaktion" samt den kontakt der findes mellem ph.d.-studerende i det jysk-fynske område. Vi har mødtes på kurser og i arbejdskredsen og holder jævnlig kontakt med hinanden. Denne kontakt er en af de positive ting ved den nye ph.d.-ordning. At den er geografisk begrænset (dog ikke helt, idet vi da også har god kontakt med ph.d.-studerende og andre forskere på den anden side af Storebælt) skyldes måske sektioneringen af ph.d.-studierne, eller rene tilfældigheder. Ordet "akse" var altså ikke ment krigerisk; jeg forsøgte at finde en betegnelse for et samarbejde som har betydet meget i denne sammenhæng og som tilfældigvis er udgået fra Odense, Århus og Aalborg.

Der kom også indvendinger mod min afvisning af at bruge sociale, kønsmæssige eller etnografiske kriterier ved udvælgelsen af samtaledeltagere. Hertil vil jeg blot sige at nogle tilgange tilstræber en repræsentativitet i forhold til sådanne "ydre" variable, mens andre tilstræber så autentiske samtaledata som muligt. De to tilgange udelukker ikke hinanden.

Vores data er ikke repræsentative i social forstand, men de repræsenterer de pågældende deltageres måde at kommunikere på i bestemte situationer. Sådanne data er velegnede til *kvalitative* studier. Hvis man vil finde ud af hvor repræsentative ens resultater er i forhold til ydre variable, så skal man have fat i andre data. Og det er vel i grunden ikke så galt!

Referencer

- Atkinson, J. Maxwell & John Heritage (eds.) 1984: *Structures of Social Action. Studies in Conversation Analysis*. Cambridge: Cambridge University Press.
- Gregersen, Frans & Inge Lise Pedersen (eds.) 1991: *The Copenhagen Study in Urban Sociolinguistics*. Vol 1 & 2. København: Reitzel.
- Gregersen, Kirsten 1992: Dansk standard for udskrifter og registrering af talesprog. 2. udgave. Tillæg til SNAK 5. årgang. København: Statens Humanistiske Forskningsråd.
- Heritage, John C. 1985: "Recent Developments in Conversation Analysis", i: *Sociolinguistics*, Vol. XV, 1.

- Ladefoged, Jakob og Diana Ilcus 1993: *Må jeg give et andet eksempel. Samtidig tale i forskellige samtale typer*. Århus: Inst. f. Lingvistisk, Aarhus Universitet, upubl. emneopgave.
- Ladefoged, Jakob 1994: *Projicering og overgangsrelevans*. Århus: Inst. f. Lingvistik, Aarhus Universitet, upubl. emneopgave.
- Linell, Per & Lennart Gustavsson 1987: *Initiativ och respons. Om dialogens dynamik, dominans och koherens*. Linköping: University of Linköping. Studies in Communication.
- Nielsen, Eva Husted 1994: - "Så går vi rundt om en uenighedsbusk!" - "Gør vi?" Århus: Inst. f. Lingvistik, Aarhus Universitet, upubl. emneopgave.
- Pedersen, Lotte Meinby 1994: *Præference i ytringspar. En analyse af præferenceforhold og opnåelse af enighed i en dagligdagssamtale på dansk*. Århus: Inst. f. Lingvistik, Aarhus Universitet, upubl. emneopgave.
- Saville-Troike, M. 1982: *The Ethnography of Communication*. Oxford: Basil Blackwell.
- Schegloff, Emanuel & Harvey Sacks 1973: "Opening up Closings" i: *Semiotica*, Vol. 8.
- Steensig, Jakob Ilum 1993: "Dansk set udefra: En analyse af overgangssignaler i to danske samtaler, set i et kontrastivt perspektiv" i Kunøe, Mette & Erik Vive Larsen: 4. *Møde om Udforskningen af Dansk Sprog. Aarhus Universitet 8.-9. oktober 1992*. Århus: Institut for Nordisk Sprog og Litteratur.
- Steensig, Jakob 1994a: "På tur i talens sprog - om syntaks og tale i samtaler" i: *NyS* nr. 18. København: Dansklærerforeningen.
- Steensig, Jakob Ilum 1994b: "Nogle overvejelser om talesprogsgrammatik" i: *SPRAU. Sprogvidenskabelige arbejdsrapporter fra Aarhus Universitet* nr. 1. Århus: Aarhus Universitet.
- Tolstrup, Anja 1994: *Rep- Re:pe: Reparatur?* Århus: Inst. f. Lingvistik, Aarhus Universitet, upubl. emneopgave.