

Mål & Mæle · 2

13. årgang 1990

Skriftsprog og talesprog

Forskellen mellem tale og skrift er ikke bare at det ene udtales og høres det andet skrives og læses. Talesprog og skriftsprog er delvis organiseret efter meget forskellige principper, afhængigt af hvordan de to slags sprog bruges. Læs om forskellen mellem tale- og skriftsprog

Side 9

Udtaleforskelle

Dansk kan tales på mange måder. Udtalen veksler fra egn til egn, fra generation til generation, fra person til person, og den enkelte kan variere sin udtale fra situation til situation. Der er foretaget en stor undersøgelse af udtalevariationen i København, og Mål og Mæle kan nu bl.a. bringe den rystende sandhed om det flade a

Side 17

Talemaskiner

Når man udveksler ordrer og beskeder med en computer, foregår det ved hjælp af skriften: Computeren og brugeren skriver til hinanden. Men der vil være mange fordele ved også at kunne kommunikere mundtligt med en datamaskine, altså give sine ordrer mundtligt og modtage sine svar som hørbar tale. På Københavns Universitet arbejder man med konstruktionen af en maskine der kan producere tale

Side 24

Sprogviden

Mål og Mæle vil i dette og en del følgende numre bringe et sæt sproglige opgaver. Det er vi blevet opfordret til af mange læsere. Men altså: Det er opgaver, ikke konkurrencer – i al fald er det ikke tanken at man skal indsende svarene, og præmier er der slet ikke! Men man kan jo udmærket arrangere konkurrencer i familien eller mellem venner og kolleger. Løsningen står på side 8 og selve opgaverne

Side 32

Sprogligheder ... 2

Talesprog 9

Sociale udtale-
variationer
i København 17

Tekst, tale og
højtælende
maskiner 24

Sprogviden 32

Det danske sprog er en møgbeskidt affære. Det overtræder sine egne forbud og regler og er fyldt til randen med overflødige ord, en mødding.

Per Højholt

SPROGLIGHeder

Sprogligheder

Denne brevkasse handler om sprogligheder. Det er spørgsmål om problemer om sprog, men det er også fine detaljer i sproget som man bliver opmærksom på, og som man vil gøre andre bekendt med. Går I rundt og tænker på sprogligheder, så send et brev om dem til Erik Hansen, Ole Tøgeby og Carsten Elbro. De vil svare på brevet hvis de kan. Ellers kender de nok nogen de kan sætte til det. Send brevet til:

G·E·C GAD
Vimmelskafte 32
1161 Kbh. K.

og mærk kuverten Sprogligheder.

? Kære Mål & Mæle

En journalist brugte i et TV-interview udtrykket "vi danskere", hvortil den interviewede, nemlig Victor Borge, bemærkede: "Det hedder da "vi danske", det ved man da!". Gør man?

Med venlig hilsen
Knud F. Franck
Frederiksberg.

! Kære Knud F. Franck

Jeg har aldrig hørt at *vi danskere* ikke skulle være lige så godt og rigtigt som *vi danske*, og jeg ved ikke hvor Victor Borge kan have den særlige idé fra. Gad vidst om han også ville foretrække *vi amerikanske* frem for *vi amerikanere*? Det kan man dårligt tænke sig.

Men det kan jo være han har tænkt på betydningsforskellen mellem de to udtryk. *Vi danskere* betegner folk med dansk pas og statsborgerskab; så at sige en rent administrativ betegnelse. *Vi danske* derimod er folk med dansk sprog og sindelag, opvokset i den danske kulturtradition – altså en lidt mere højtidelig og inderlig betegnelse.

EH

? Kære Mål & Mæle

Et spørgsmål har "uroat" mig meget længe. Hvis man er lunefuld, så er man månesyg (via plattysk fra latin *luna*), moody og ikke til at være i nærheden af. Hvis man har lune (også fra latin *luna* over plattysk), så er man en morsom fyr, nærmest lun, som er et fællesnordisk ord, beslægtet med *lue* og *lys*. Det er i familie med svenske *lugn*, der som bekendt på dansk betyder 'rolig' (undtagen i vendinger som *ligge lunt i svinget*), der som bekendt på svensk betyder "morsom", for ikke at sige "lun".

Hvad er meningen, hvordan er sammenhængen, og hvorfor betyder *urolog* urolig på både svensk og dansk? Jeg er forvirret.

Gorm Rasmussen
Dannemare

! Kære Gorm Rasmussen

Et *lune* ('en øjeblikkelig sindsstemning') har sproghistorisk andre rødder end tillægsordet *lun* ('varm'). Så lad os se på dette *lune* og *lunefuld* først.

Et *lune* kan ganske rigtigt føres tilbage til latin *luna* ('måne'). Ordet er lånt via tysk i betydningen 'øjeblikkeligt, omskifteligt sind' selvfølgelig ikke i betydningen 'måne', for den har vi jo et fællesnordisk ord for. Hvis man er fuld af omskiftelige sindsstemninger, kan man derfor udmærket kaldes *lunefuld*. Men *lune* bruges desuden om en særlig, stilfærdig humor, der gerne udnytter underdrivelser. *Selv da Sørensen mødte sin overmand, fornægtede hans lune sig ikke. Lille Helle satte pris på bedstemors landlige lune.* Man kan måske sige, at den venlige dobbeltbundethed, blinket i øjet, det skælmiske i at have *lune* viser tilbage til det 'øjeblikkelige, omskiftelige sind', men rigtignok i positiv betydning.

Så er der tillægsordet *lun*, der er det samme som svensk *lugn* ('rolig'). Det er afledt af det nu forældede navneord et *lun* med grundbetydningen 'et sted hvor vandets overflade er blank', d.v.s. et sted hvor der følgelig er stille, roligt og lunt. Udviklingen i svensk har så lagt vægten på betydningen 'stille og roligt'; hvorimod udviklingen i dansk i højere grad har fokuseret på 'varme'.

Anden del af spørgsmålet drejer sig om de 'falske venner' *rolig* på dansk og *rolig* ('morsom') på svensk. Begge ord er afledt af det samme ord *ro*, hvis oprindelige betydning er 'hvile'. Betydningen af det danske *rolig* er således helt ligefrem 'i hvile, stille, fredeligt'. Det svenske *rolig* (og ud-sagnsordet *roa sig*) er derimod afledt af en anden betydning af *ro*, nemlig 'morskab, tidsfordriv', d.v.s. hvad man foretager sig under hvilen f.eks. fra arbejdet. Denne anden betydning af *ro* har vi også haft i dansk, men den bruges ikke mere.

Det påfaldende består altså i, at svensk har afledt *rolig* af én betydning

af *ro*, hvorimod dansk har afledt *rolig* af en anden. Der er ikke noget mærkeligt i, at begge sprog har afledt *uro* (og heraf *urolog* og svensk *oroa*, 'gøre urolig') af den samme betydning af *ro*, nemlig 'hvile'.

CE

? Kære Mål & Mæle.

Hvornår er man kendt?

Hvis en person er kendt, lægger vi tryk på hans efternavn (1. stavelse) i stedet for lige tryk på fornavn og efternavn. F.eks. *Peter Olesen* – hvem er han? Men *Peter Olesen*, det er jo ham fra TV-avisen

Det virker faktisk sådan, at man kan afprøve, hvor kendt en person er – eller man selv synes man er. Ringer man f.eks. til en kendt person i et firma, hvor man selv er kendt, siger man: "Goddag, det er *Peter Hansen*, må jeg tale med *Ole Petersen*. Men er det tale om to fremmede, siger man: Mit navn er *Peter Hansen*, jeg vil gerne tale med *Ole Petersen*".

Ringer man til Mål & Mæle for at tale med den landskendte Erik Hansen, siger man: Goddag, det er *Erik Hansen*, må jeg tale med *Erik Hansen*". osv. osv.

Jeg nævnte dette for nogen tid siden for en sprogmand. Han syntes ikke, det var værd at bemærke, men almindelige mennesker siger, at de aldrig før har tænkt over det, men at det er ganske pudsigt.

Med venlig hilsen
Marianne Thillemann

! Kære Marianne Thillemann

Jeg tror din iagttagelse er ganske rigtig, og hvis din regel ikke er fuldstændig fast, så er der i al fald tale om en meget klar tendens.

Jeg har kun en enkelt bemærkning. Udtalen med tryksvagt fornavn og trykstærkt efternavn (Jens *Petersen*) forudsætter ikke nødvendigvis at personen egentlig er kendt. Hvis jeg fortæller dig om min nabo – som du ikke kender – så siger jeg *Henrik Lund* første gang jeg omtaler ham. Derefter går jeg over til *Henrik Lund*, altså med tryksvagt fornavn og trykstærkt efternavn. Og det hænger sammen med at nu ved du hvem han er – uden at *Henrik Lund* behøver at være berømt og omtalt i *Se & Hør*.

Det hele minder meget om betydningen af bestemt og ubestemt artikel. Jeg siger fx *et hus* første gang, for jeg regner ikke med at du ved hvad jeg taler om; det svarer til navne med tryk på både for- og efternavn. Men i det følgende siger jeg *huset*, for nu er du med på hvad det drejer sig om; det svarer til navn med tryksvagt fornavn og trykstærkt efternavn.

Hvad var det for en sur sprogmand der ikke syntes du havde gjort en interessant iagttagelse?

EH

? Kære professor Erik Hansen. Tak for den lille artikel "Ret og rime-
ligt" i Mål & Mæle fra dec. 89.

Første gang jeg stødte på "rigtig" som forstærkende ord, var i et brev som jeg fik i sommeren 1984.

"Rimelig" kom vel et par år senere, og irriterer mig stadig. Hvis det virkelig er kommet for at blive i den nye betydning, altså forstærkende, vil jeg sikkert vænne mig til det i tidens fylde, ligesom det er gået med ret og temmelig, som du nævner.

Hvad jeg egentlig skriver om til dig, er den mere og mere almindelige brug af ordet "utrolig". Har du lagt mærke til det? Det bruges ustandseligt som

forstærkende ord nu – jeg hørte det således brugt første gang i november 1989 mundligt. Det har fået en hurtigere og større udbredelse – så vidt jeg kan se – end ordet "rimelig", og det hænger vel sammen med, at det ikke i forvejen har haft anden betydning end netop den forstærkende?

Men – hvorfor bliver et ord pludselig moderne?

Med venlig hilsen
Else Varnild
Odense C.

! Kære Else Varnild.

Ja, det er rigtigt at *utrolig* som forstærker har grebet voldsomt om sig de sidste to-tre år. Der er folk som har været på ferie i Jylland *utrolig* mange gange, man bliver *utrolig* glad for reaktionerne, det er *utroligt* som blomsterne vælder frem og Lise taler *utrolig* godt fransk.

Udtrykket har for længst passeret klichétærskelen, og der er gode argumenter for at vise lidt tilbageholdenhed med brugen af ordet. Det er faktisk meget irriterende!

Hvorfor et ord eller et udtryk pludselig bliver moderne, kan man som regel ikke forklare. Det hænder at en kendt person, fx en politiker, kommer til at bruge et udtryk som slår an, og så kan det normalt spores. Som da socialdemokraten Per Hækkerup i 1967 fandt på ordet *stokkemetode*, eller da dronningen i sin nytårstale 1984 lancerede tillægsordet *dumsmart*. Men i reglen har man som sagt ingen forklaring.

Når det drejer sig om forstærkerord af typen *utrolig*, er der det særlige at der altid synes at have været et behov for et stort udvalg. Prøv engang at slå op i en synonymordbog, fx Politikens

Forlags under *særdeles*. Det er en af ordbogens længste artikler, 27 linjer med forstærkende synonymmer: *betydelig, drabelig, enorm, morderlig, overordentlig, rigtig, såre* osv. Der er virkelig noget at vælge imellem, så det er måske ikke sært at der kommer nye ord til – området ser ud til at være meget produktivt. Det morsomme er at *utrolig* ikke er med i listen (men dog andetsteds kan findes som synonym til *usandsynlig*). Sidste udgave af Politikens synonymordbog er fra 1988, så det er først fra omkring den tid at *utrolig* har indledt sin hængende sejrsgang.

EH

? Kære Mål & Mæle

Jeg skal være den sidste til at bekende mig til sprogrensernes rækker.

Alligevel kan jeg af og til blive lidt vemodig, når gode danske ord bukker under for udenlandske, der ikke er spor bedre eller mere dækkende.

Hvorfor går alle rundt og siger *jaloux* og *ambitiøs*, når vi dog har de danske ord *skinsyg* og *ærgerrig*? Ordet *jaloux* har jo desuden en for danskere sær stavemåde. Kan I sige noget generelt om, hvilke danske ord der på denne måde 'dør' og hvilke ikke? Og er ordene *skinsyg* og *ærgerrig* faktisk døde, sådan som jeg påstår?

De samme betragtninger kan man vel gøre med hensyn til ordene *generation* kontra *slægtled* og med hensyn til *melankolsk* kontra *vemodig*.

Inden for EDB-verdenen taler man om, at *datamater* (undskyld: computer) er understøttet med *hard-disk*, hvilket på dansk hedder *hårdpladelager*. I den sammenhæng kunne man måske (forgæves?) slå et slag for, at ordene *hardware* og *software* erstattes af de helt ensbetydende danske ord *maski-*

nel og *programmel*? Jeg er ganske klar over, at EDB-verdenen er specielt inficeret (*besmittet*?) med engelske ord, der da af og til ikke har nogen danske modstykker.

Med venlig hilsen
Jens Carstensen

! Kære Jens Carstensen

Dit brev indeholder så gode ideer og så svære spørgsmål at vi har besluttet snarest at bringe en hel artikel om ordenes liv og død og om dansk og fremmed i ordforrådet. Derfor får du kun nogle spredte bemærkninger her.

Retskrivningsordbogen (1986) indeholder ca. 60.000 ord. Det er ikke hele det danske ordforråd, men en pæn del af det. Man regner med at kun ca. 2.000 af disse ord er ægte danske, i den forstand at de ikke er lånt fra et andet sprog. Det viser i hvilket omfang dansk i tidens løb har ladet sig berige af andre sprog.

Vi er derfor vant til at omgås ord fra andre sprog, og vores oplevelse af dansk forstyrres ikke af at vi har så mange ord udefra. I mange tilfælde tror vi at et ord er pæredansk, men opdager ved at kigge i ordbogen, at det forholder sig anderledes.

Du nævnte fx de "danske ord" *skinsyg* og *ærgerrig*. Tjæ, *ærgerrig* er i hvert fald rent tysk, *skinsyg* måske også, men det er lidt usikkert. Også *vemodig*, som du oplever som purt dansk, er klokkerent tysk, og det er *slægt* i *slægtled* også.

Men det er da rigtigt at sådanne gamle tyske ord ikke virker nær så fremmede som franske, latinske og græske: *jaloux*, *ambitiøs*, *generation*, *melankolsk*. Endnu mere fremmede virker nok engelske ord, bl.a. fordi vi som hovedregel prøver at sluge dem råt, dvs. med original stavemåde,

udtale og bøjning: *harddisk, hardware, software*. Derfor virker, som du skriver, *maskinel* og *programmel* mere hjemlige, selv om det faktisk er latin og græsk med en endelse der nærmest er fransk!

Men som sagt vi vender tilbage til sagen.

EH

? Kære Mål & Måle.

Jeg blev meget glad, da jeg læste på forsiden af M & M 1/1989, at problemet med ordet "så" som konjunktion ville blive behandlet.

Jeg kan fra den pædagogiske front meddele, at dette lille ord skaber uanede problemer i undervisningen af udlændinge. Som en af mine udenlandske elever en dag sagde:

"Jeg bruger aldrig *så*, for jeg kan ikke finde ud af, hvornår man skal bruge det på den ene måde og hvornår på den anden måde!"

I M & M 1/89 havde Tue Eriksen fra Århus problemer med sætningen: "Det er meget varmt, så du behøver ikke at tage overtøj på!" Et klassisk sprogligt problem for udlændinge.

Med stor interesse begyndte jeg så at læse svaret. Det hele lød meget rigtigt – uden egentlig at være nogen forklaring.

Jeg kom hurtigt til at tænke på en passage fra Erik Hansens udmærkede bog *Da lo hun så hjerteligt*, hvor der står: "Den gode tekst er så at sige usynlig, den stiller sig ikke uforstyrrende i vejen mellem læser og information!"

I svaret til Tue Eriksen fik vi (læsere) en masse forskellige, rigtige men uvedkommende ting at vide om forskellige sætningers kombinationsmuligheder. Disse oplysninger virkede

mere "uforstyrrende" end forklarende.

Stor var min for undring imidlertid, da jeg læste mig til, at ordet "nøglen" skulle være et biord. Her måtte der være noget galt.

Svaret fortsættes med at vi rigtigt nok belæres om, at "så" osse kan udtrykke *tid alene* og ikke årsag; men..... det eksempel, vi fik som belæg for dette, var en sætning, der helt klart taler om en årsags-sammenhæng, nemlig: "fordi.....så" – sammenhængen: "(fordi) han kom, så gik jeg ikke".

Dette fik mig til at tænke på endnu en passage fra Erik Hansens bog, nemlig:

"Det dårlige sprog er som en snavset rude med støbefejl; det slører, forstyrrer og forvrænger" (s.64).

Hvad er det, der skal sløres i dette svar til Tue Eriksen? Det skulle vel ikke være det, at forfatteren til svaret selv har problemer med ordet "så"?

Selv efter 5. gennemlæsning er det ikke muligt at finde meget formidlende (eller formildende) i svaret. Svaret er elendigt på en nærmest monumental måde. Det er hjælpeløst og sjusket lige fra tegnsætning til disposition.

Jeg ved ikke, hvordan Tue har det med svaret. Jeg kan kun mene: OM IGEN!

Jeg glæder mig til næste nummer af Mål & Mæle, hvor jeg ser frem til en nøgtern forklaring på problemet med ordet "så" som henholdsvis hovedsætnings- og bisætningskonjunktion.

Med venlig hilsen
Henrik Krogager
Århus

! Kære Henrik Krogager.
Tak for et barsk brev. Der var i den trykte udgave gået kludder i nøgleaf-

snittet. Det skulle lyde således (som der stod i manuskriptet):

”Man kan også udtrykke virkningen i en helsætning, med biordet *derfor*, og hvis *derfor* indleder sætningen kommer udsagnsled før grundled (inversion): *Han har lavet nøglen, derfor kan han godt komme ind*; og årsagen kan udtrykkes i en efterstillet helsætning med biordet *nemlig*: *Han kan godt komme ind, han har nemlig lavet nøglen*. En helsætning kan også indeholde og indledes af biordet *så* men det udtrykker alene tid og ikke årsag: *Han kom, og så gik jeg*. ”

Det er altså *nemlig* der er et biord, og ikke *nøglen*. Nu kommer så en forklaring på problemet med ordet *så* som henholdsvis gradsbiord, forholdsbiord, tillægsagtigt henvisningsord, helsætningsbindeord, ledsætningsbindeord og udråbsord.

1. Ordet *så* kan optræde som påpegende og henvisende gradsbiord der lægger sig til tillægsord eller biord og henviser enten til højre (kataforisk) eller venstre (anatorisk); her betyder det 'i den grad': *Det er ikke så varmt i dag som i går; han er så syg at han ikke kan komme; så dårligt havde jeg ikke troet det ville gå*. Den påpegende betydning er gået tabt i den absolutte brug, dvs. uden noget at henvise til; *så* kommer til at betyde 'meget': *Hun er så dum så dum; det var så ulykkeligt*.

2. Det kan også være et (venstre)henvisende forholdsbiord der lægger sig til udsagnsledet og betyder 'derpå', 'i det tilfælde', 'under disse omstændigheder': *bliv ulykkelig først og græd så; hvis han kommer, hvad skal jeg så sige?; han kom og så gik jeg; han må være kommet ind her, og han må så være gået derhen*. Som forholdsbiord kræver *så* omvendt ordstilling (dvs. grundled efter udsagnsled) hvis

det i en helsætning står først, dvs. lige efter et evt. *og*.

3. Ordet *så* kan (når det rimer på *lå*) være et henvisningsord der kan minde om et tillægsord og som betyder 'således', 'den' eller 'det': *...om jeg så må sige, ...i så henseende*.

4. *Så* kan som bindeord indlede en helsætning (hovedsætning) uden at der kommer omvendt ordstilling – ligesom der heller ikke kommer omvendt ordstilling ved bindeordet *og*. Ordet betyder da 'og følgelig' (læg mærke til forskellen på 'som logisk følge' og 'som virkning af en årsag'). *Hun var syg, så hun kunne ikke komme; de opdagede forklædningen, så den gik ikke*. Læg mærke til forskellen på sætninger med biordet der i betydningen 'derpå' eller 'under disse omstændigheder' indleder en helsætning med omvendt ordstilling: *Han stod op, og så gik han på arbejde*, og bindeordet der i betydningen 'følgelig' indleder en helsætning uden omvendt ordstilling: *Han stod op, så han var ikke syg*.

5. Ordet *så* kan som bindeord også indlede en ledsætning (bisætning), til venstre for et evt. *at*, og have ledsætningsordstilling, dvs. billedet til venstre for udsagnsledet. Ordet betyder her: 'med den konsekvens at': *de havde arrangeret det så (at) han ikke skulle komme*. Som ledsætningsbindeord kan *så* blive kædet sammen med ordet *ikke* således at begge ord tilsammen optræder som ét ord på den plads som *så* ellers havde: *de havde arrangeret det så ikke han skulle komme*. Som ledsætningsbindeord kan *så* (*at*) vise tilbage til et højrehenvisende gradsbiord: *Han var så fuld så han ikke kunne gå lige*.

6. Endelig kan *så* være et udråbsord: *Så, så, tag det nu roligt!*

Sætningen *Fordi han kom, så gik jeg*

ikke er et eksempel på så 2, det venstrehenvisende forholdsbiord; årsagsbetydningen ligger i ordet *fordi* og *så* henviser blot til denne omstændighed. Hvis der ikke er noget *fordi*, er der heller ikke nogen årsagsbetydning: *Han kom og så gik jeg* betyder snarere 'jeg gik da han var kommet', end 'jeg gik fordi han kom'. Sætningen *Han kom, så gik jeg ikke* bliver næppe sagt af nogen.

Var det du så, så så, så at sige, nøgternt så det var så som så med så barske ord?

OT

Løsninger

1. Den mest almindelige udtale i dansk er muligvis [stik], men på engelsk er det [stæjk], så det er den man trods alt må anbefale. Iøvrigt er det et oldnordisk ord, *steikr*, som vikingerne bragte med til England, så den rigtige udtale stemmer også med oldnordisk!

2. Klokkeslæt – hefte/hæfte – delle – egypter/ægypter. Et kort æ skrives i de fleste tilfælde som e: *vende, hest, sted*. Hvis et ord er beslægtet med eller er en bøjningsform af et ord, der indeholder a eller å, så skrives et kort æ som æ: *bånd/bændel, and/ænder. -slæt i klokkeslæt, udslæt og islæt* er beslægtet med *slå*, derfor æ; *hefte/hæfte* er ikke beslægtet med nogen a-ord i dansk, derfor skriver de fleste *hefte*, og den form blev indført som korrekt ved siden af *hæfte* i 1986. Skrivemåden *dælle*, der ikke har nogen a-slægtninge og som ingen brugte, blev afskaffet i 1986. Samtidig indførtes *egypter, egyptisk* og *Egypten* som mulige former; påvirkningen fra engelsk havde gjort æ-formerne meget sjældne.

3. *Uden om* i b og c. Reglen er at hvis det er et forhold-sord med en styrelse, så skrives det i to ord, ellers i et ord. Altså *uden om skoven, ...bog, som...uden om*. Se retskrivningsordbogen side 545-46.

4. Alle tre sætninger har korrekt grammatisk komma, også c: Her har vi en såkaldt sætningsknode, dvs. en konstruktion, hvor et led fra en bisætning ([at] vi skal vente at se...igen) er anbragt i spidsen af hovedsætningen: *Ham*. I sådanne tilfælde behøver man ikke at sætte komma foran bisætningen. Se retskrivningsordbogen side 584-85.

Mål & Mæle

Redaktion:

Erik Hansen, Ole Tøgeby, Carsten Elbro
Tegninger: Anne M. Nielsen

Ekspedition: G.E.C. Gad, Strøget Dansk og Udenlandsk Boghandel AS, Vimmelskafte 32, 1161 København K.
Telefon 33 15 05 58 Giro 1 00 00 55

Tryk: Hans Jørgensens Bogtrykkeri
København.

Mål og Mæle udkommer 4 gange om året, og abonnementsprisen er kr. 128,00 pr. årgang. Man kan tegne abonnement ved at skrive eller ringe til Gads Boghandel. Hertil retter man også henvendelse om adresseforandring eller fejl ved bladets levering. Eftertryk af tekst og illustrationer er tilladt når kil-den angives.

5. Det hedder *billede* og *hoved* i overensstemmelse med de to ords historie: gammeldansk *bilæt* og *houæth*. Efter samme princip burde det have været *hundred*, men da *tusind(e)* ganske ubistorisk har fået en form med -e, er formen *hundrede* indført ved side af *hundred*.

6. *Speaker* er formanden for Underhuset; en *choker* er et halstørklæde, en *krave* eller en *flip*; en *speeder* er en fartbølge; *smoking* betyder rygning. Men en *spencer* er en spencer!

7. *Gro og grøn*.

8. Vrvølet findes i b: *Merudgifter* kan ikke være noget, man planlægger! Hvis man finder vrvøl i a, er man en stor pedant. Det, man *tager af*, kan naturligvis godt være tøj, men ikke når det er vind og regn; så betyder *tage af* det samme som 'mindskes'.

9. *Som der* er indiskutabelt korrekt i a. Reglen er, at en bisætning er korrekt med *som der*, hvis der et også findes i den tilsvarende hovedsætning. Det gør det ikke i fx c: *det samme (argument) blev brugt sidste gang*, og derfor er *som der* ukorrekt her. Og tilsvarende i b. Men i a er den tilsvarende hovedsætning *der kan ikke udsættes noget på den slags mennesker*, altså er *som der* OK.

10. Kunsten at finde skjulte vandårer med en pilekvist. Græsk *rhabdos* betyder 'kvist', og *manteia* 'spådom'.

Talesprog

Her er et uddrag fra Højlunds Forsamlingshus i TV 2 (24.10.1989). Er det godt eller dårligt sprog, manden taler?

Mand: ...*det der er sket frem til i dag er rent faktisk / at den u-landsbistand vi har givet de sidste 20 år / har været med til at kreere nogle hvide elefanter mange steder henne / det er ikke det det handler om / det det handler om det er at komme til at kreere et aktivt erhvervsliv / øh så er der jo nogle der siger jamen har I i industriet også ikke også været med til at / lave nogle af de der hvide elefanter jo det har vi måske nok / men vi har været med til at / vi har altså spillet på nogle betingelser som der har været andre der tidligere har / ligesom sat reglerne for / men jeg føler i og for sig / at / vi er ved at komme til at leve i en pragmatisk verden / hvor at / vores interesser for erhvervslivet u-landsfolkenes interesser for at de kan mødes / det vil altså sige...*

Højlund (afbryder): *hvad mente hvad mente du det du sagde før hvide elefanter altså projekter der ikke blev til videre og så videre*

Mand: *ja det mente jeg faktisk*

I første omgang viser eksemplet os, hvordan talesprog ser ud, når det bliver skrevet ned. Der er hverken kommaer eller punktummer, men nok en slags pauser, markeret med skråstreg.

Men også på andre måder ser det mærkeligt ud på tryk. Der er *øh*'er,

sprogfejl, selvrettelser og sætninger, der er gået helt i stykker, uden at den talende har forsøgt at gøre det godt igen, og der er en del af den slags ord, som man plejer at kalde fyldeord: *faktisk, altså, ligesom, i og for sig*.

Er det nu et eksempel på særlig dårligt talesprog, vi har her? Nej, absolut ikke. Det særlige ved det er kun, at det er blevet skrevet ned og senere trykt. Det er et eksempel på ganske almindeligt talesprog.

Tale og skrift

Talesprog og skriftsprog er naturligvis ikke to forskellige sprog, sådan som dansk og engelsk er forskellige sprog. Vi taler jo også om dansk talesprog og dansk skriftsprog, og det viser, at begge dele er dansk. Dansk har bare, som de fleste andre sprog, to hovedfremtrædelsesformer: en lydlig og en visuel. Vi kan altså opfatte sproget med øret eller med øjet.

Det giver en meget stor forskel mht. det rent fysiske, men det meste er fælles for talesprog og skriftsprog, fx ordforråd og grammatik. I sprogforskningen har man ældgammel tradition for at studere skriftsproget, når man vil vide noget om sproget. Det hænger naturligvis sammen med, at det har været det letteste, og næsten det eneste mulige, lige indtil opfindelsen af båndoptageren for få årtier siden.

Naturligvis har man vidst lidt om talesproget før, men den viden, forskere og lægfolk har haft om talesproget, har ikke været tilfredsstillende, bl.a. ved at

fremstille det talte sprog som en primitiv og ufuldkommen form af det egentlige og rigtige sprog, nemlig skriften. Det er der to grunde til.

For det første havde man af de omtalte tekniske årsager ikke gode muligheder for at iagttage og analysere de mange midler, som den talende benytter for at gøre ytringerne klare og letforståelige, fx rytme, tonegang, pauser, tempo.

For det andet består dette at lære at skrive korrekt for en stor del i at aflægge nogle af de vaner, man har fra talesproget, fx særlige ord og vendinger, selvrettelser, ufuldstændige sætninger, specielle småord. Set fra skriftens synspunkt bliver de særlige talesprogstræk derfor let opfattet som fejl, simpelt hen.

Talens fysiske betingelser

Den moderne sprogsforskning betragter naturligvis ikke talen som en sjusket udgave af skriftsproget, og man arbejder i mange lande alvorligt på at udforske talens grammatik som et system, der nok har det allermeste til fælles med skriften, men alligevel er selvstændigt og har sine egne love og regler, således som skriften har sine.

De træk, der er karakteristiske for talesproget, kan alle afledes af talens fysiske betingelser. Talen frembringes i reglen spontant, altså uden planlægning og kladde; og tilhøreren forventes at opfatte talen i første forsøg – han kan ikke som ved læsning afkode teksten i sit eget tempo og genlæse det, han ikke fik fat i med det samme; det er altid irriterende og småpinligt at bede om gentagelser (*hvabehar?*).

Desuden er tale næsten altid samtale mellem to eller flere parter. Delta-gerne har derfor en meget klar bevidsthed om hinanden, om det rum, som

samtalen foregår i, og det tidspunkt, hvor kommunikationen finder sted.

Vores oversigt over de mest karakteristiske talesprogstræk vil vi da indrette efter disse betingelser: (1) Spontanitet og tempo, (2) Samtaledeltagernes bevidsthed om hinanden, (3) Samtaledeltagernes fælles bevidsthed om rum og tid.

I den følgende gennemgang af de vigtigste træk ved talesproget er eksemplerne næsten alle taget fra nogle moderne skuespil. Man kunne naturligvis også have hentet eksemplerne fra autentisk, opnoteret talesprog, men fordelene ved at tage dem fra skuespil er, at man så lettere kan nøjes med ét træk ad gangen. I de autentiske tekster får man det meste i hovedet på én gang, og det kan virke meget uoverskueligt.

Korte enkle sætninger

Talesprogets tempo spiller en stor rolle for dets indretning: Den talende kan normalt ikke planlægge meget store og indviklede tekstsammenhænge, og tilhøreren kan ikke overskue andet end det ret enkle. Dette viser sig på alle områder. Talesproget bruger *korte sætninger*:

Jeg måtte tale med fru Møller – men der var ingen lejlighed – så hørte jeg, at hun ville komme på skolen om aftenen for at ordne nogle bøger til skolebiblioteket – det havde hun under sig – det ligger helt oppe øverst – det er loftet, der er lavet om – skolebetjenten – Rasmussen hedder han – han bor nede i kælderen – han er gift og har en datter – der var lys nede hos Rasmussen, da jeg kom – (ASH 38)

At sætninger kan være korte, hænger sammen med, at talesproget ikke bru-

ger mange bisætninger. I normalt skriftsprog kan man fx udtrykke sig sådan:

Jeg gik rundt uden arbejde i de år, hvor arbejdsløsheden var allerværst. Når folk spurgte mig, hvad jeg kunne, måtte jeg svare, at jeg ikke kunne noget. Men en dag var jeg så heldig at få en plads som elevatorfører i en stor forretning med pelse, kjoler og den slags.

Men i talesprogsudgave ser det sådan ud:

Så gik jeg et par år og spadserede. Det var lige i de år, hvor arbejdsløsheden var allerværst. Og hva' ku' jeg! når folk spurgte. Ikke en stift. Men så endelig, så gik der jo hul på godteposen, og jeg fik plads som elevatorhusar, det var i en stor butik med kjoler og pelse og den slags. Men de sku' aldrig ha' ta'et mig, aldrig!, jeg havde uheldet med under fodsålerne – otte dage efter jeg var begyndt, så brændte hele skidtet! (BM 10)

Talesprogets sætninger er ikke alene korte, de er også opbygget enkelt. Et af de vigtigste træk er, at der er meget kort hen til sætningens udsagnsled. Det er først når udsagnsleddet er faldet, at man egentlig kan orientere sig i sætningen, derfor står der i talesproget ét kort led foran udsagnsleddet, i det følgende eksempel *det, hun, hvad, jeg* osv.:

– *Det må jeg nok sige ... hun holder sig fantastisk ... hvor gammel er din mor egentlig?*

– *Mor? Hun er snart halvtreds.*

– *Det er umuligt! Hun ser ud, som om hun var tredve.*

– *Nej, det er mig! ... jeg er snart tredve.* (DB 53)

Og hvis en sætning er kommet til at begynde med et længere led (det behøver ikke at være ret langt!), så markeres det med et *så, der, hun, den* eller lignende ord, at nu er vi færdige med indledningsleddet og klar til udsagnsleddet:

År du ve' – når man køber sådden en butik, så blir den altid fremstillet som bedre end den er. Varelageret blir vurderet for højt, og sæl' om det fremgår af bøgerne, at omsætningen har været dalende – nåja så er det jo kun, fordi ejeren har været syg, forstår du. (FV 53)

Her er det de to *så'er* der markerer grænsen mellem indledning og resten af sætningen. Men metoden bruges også ved meget kortere indledninger: *I de første dae, da må du ha' et lille hotelværelse. (BM 109) Ve De, hvodden det gik mig dengang på sessionen. Ham lige foran, han blev kasseret ved det, at han var svajrygget, og ham lige efter – han trak frinummer. (BM 11)*

Fokus

I det talte sprog bruges der mange *kløvninger*, dvs. konstruktioner af typen *det var Peter, der gjorde det* (jf. *Peter gjorde det*), *det, jeg mener, er, at...* (jf. *jeg mener at...*). Det er let, men meget pladskrævende at forklare kløvningens funktion og betydning, så vi vil her nøjes med at sige, at kløvningen fremhæver eller fokuserer på det led, der kommer efter *det...er/var*; i eksemplerne ovenfor altså *Peter* og *at...* Kløvninger bruges også i skriftsprog, men talesproget har nok særlig meget brug for fremhævelse og pointering. Vi finder eksempler på det i Højlunds Forsamlingshus: *det er ikke det det handler om/ det det handler om det er at komme til at kreere et aktivt erhvervsliv.*

Væk med det overflødige

Et andet væsentligt træk ved talesproget er *økonomiprincippet*, dvs. en teknik, der udelader ord, som er ganske overflødige på grund af sammenhængen:

- Jeg tror, det ringede.
- Øh – ringede!?
- På entrédøren.
- Umuligt! på denne tid af natten! (FV 89)

Fire replikker, men kun den første har form som en sætning; de øvrige kan let suppleres op, så de også bliver fulde sætninger, men pointen er netop, at man ikke gør det i talesproget:

- Hvad laver du?
- Sover.
- Hvad er der, skat?
- Ingenting.
- Er du fuld?
- Ja.
- Meget?
- Åh, lidt (W 27)



”Hvad laver du?”

En speciel økonomiseringsmulighed har talesproget i *udråbsordene*, fx *av*, *nå*, *tja*, *øv*, *bvadr*, *hovsa*, *waww*, *bøh*. Det er ord uden grammatik og uden fastlagt betydning, men de kan være meget udtryksfulde og sige meget på meget lille plads:

- *Sku’ vi ikke lavær og snakke om ham?*
- *Det ka’ vi godt [tænder for shaveren igen] Jeg var ellers på vej til byen for at drikke mig fuld!*
- *Hvorfor det?*
- *Tjah! Hvorfor ikke... Av for helvede..*
- *Hva’ er der i vejen?*
- *Der er overgang i det møg! [går ind i soveværelset og tager sine bukser på].*
- *Er de ikke nye?*
- *Joe! Ka’ du li’ dem?*
- *Har du råd til det?*
- *Åh joe... det går... (BLAP 104)*

Udeladelser i forhold til skriftsprogets færdige sætninger skal altså ikke betragtes som uheld eller sjusk, men som et led i tydeliggørelsen af talen. Og udråbsord er ikke nødvendigvis udtryk for sproglig armod.

Men naturligvis sker der uheld, når der tales. Der er sætninger, der aldrig bliver færdige:

- *Han svarer ikke... han må være rejst uden mig.*
- *Anne, du må spise noget...*
- *Jamen han lovede... Han svarer ikke!*
- *Var det ikke først på onsdag?*
- *Han er rejst... han er...*
- *Ja, det er han vel. (BLAP 145)*

Den talende kan afbryde og rette sig selv:

- *Et glas til?*
- *Næe tak, jeg er i bil... Jeg ved ikke, hvordan jeg skal... Tilhører De en... en... en trosretning? En speciel...? (BLAP 93)*

vi er ved at komme til at leve i en pragmatisk verden / hvor at / vores interesser for erhvervslivet U-landsfolkenes interesser for at de kan mødes / det vil sige (HF)

Læg også mærke til forbindelsen *hvor at*. Man kan diskutere, om det er en sprogfejl, men det er i al fald næsten sikkert, at den interviewede ikke ville have skrevet *hvor at*!

Og endelig kan en sætning gå i fisk for den talende, så han måske kommer til at sige noget vrøvl, som han måske aldrig ville kunne skrive: *Der er stadig resterende billetter til koncerten* (DR maj 1986). *Dagens længde er blevet 10 timer og 46 minutter længere* (DR maj 1988). *I løbet af dagen vil der brede sig et skydække ud over landet, som efterhånden vil gå i opløsning* (DR maj 1989). *DSB blev idømt til at betale en bøde* (DR okt. 1989).

Repliktennis

Prøv at læse følgende dialog opmærksomt:

- *Hvad vil du have?*
- *Jeg vil gerne have en sodavand.*
- *Vi har øl.*
- *Jeg kan godt lide øl, men jeg vil hellere have en sodavand.*
- *Det er det bedste når man er i bil.*
- *Jeg er ikke i bil. Der er kun et par hundrede meter herhen.*
- *Der er mindst 2 km.*
- *Det er ikke ret meget.*

Ikke sandt, man forstår det hele, men det er, som om de to personer slet ikke er i kontakt med hinanden, men bare står og siger replikker ud i luften. Det kommer af, at teksten mangler alle de *småord*, hvis funktion det er, hele tiden at sætte replikkerne i forhold til hinanden. Så snart de kommer med, virker det meget mere som en naturlig dialog – de talende har fået kontakt med hinanden:

- *Hvad vil du have?*
- *Jeg vil faktisk gerne have en sodavand.*
- *Vi har ellers øl.*
- *Jeg kan skam godt lide øl, men jeg vil nu hellere have en sodavand.*
- *Det er også bedst, når man er i bil.*
- *Jer er da ikke i bil. Der er jo kun et par hundrede meter herhen.*
- *Der er altså mindst 2 km.*
- *Det er da heller ikke ret meget.*

Disse småord er meget karakteristiske for samtaler:

- *Hvorfor kører du ikke med sikkerhedssele?*
- *Med hvad?*
- *Med sikkerhedssele. Har du ikke set, man kan få sådan nogen nu...til biler?*
- *Jah...men det er sgu da alt for besværligt... Så kan man jo hverken lave det ene eller det andet.*
- *Hvad skal du da lave, andet end at køre?*
- *Det ved man da aldrig. Vel, Ida?*
- *Osse mens I kører?*
- *Gud nej, det har vi forresten aldrig prøvet før. Det må vi prøve!* (DB 136)

Det er også her, talesprogets karakteristiske pauseudfylder *øh* kommer ind. Når man udveksler replikker, er en pause som regel et signal om, at man er færdig og overlader ordet til den, man taler med. Derfor kan det være farligt at holde pause, mens man tænker sig om, og det er grunden til, at man siger *øh*. Det betyder altså "jeg har stadig ordet".

Rum, tid og person

Deltagerne i en samtale har altid en fælles viden om den talesituation, de netop befinder sig i. Først og fremmest

kender de det rum eller det sted, hvor de befinder sig, og de kender det tidspunkt, hvor samtalen foregår. De kan derfor bruge stedsbetegnelser som *her, derhenne, oppe, højre, venstre, foran, bagved* osv., idet ordene tager betydning efter situationen. På samme måde med tidsbetegnelserne: *nu, for en time siden, i forgårs, om lidt*, osv. Dem må man tit undgå i skrift, hvor parterne jo i reglen ikke befinder sig i samme situation: Man kan ikke skrive *i overmorgen*, det må være *den 12. marts 1990*.

Men talesproget er fuldt af disse ord, som forudsætter fælles situation. Se fx på stedsbetegnelserne her:

- *Jeg var hér ude og... hænge tøj til tørre... Småvask du-vé... og så ser jeg pludselig en mand dérovre på en af altanerne.*
- *Hva' for en?*
- *Den dér.*
- *Den der med gyngestolen?*
- *Næ, den til højre. (FV 157)*

Til disse situationsord hører naturligvis også *jeg, du, han/hun, vi, I, de*. Læg mærke til, hvor ubesværet de antager betydning i en naturlig samtale:

- *Davs!... Nå... du er her?*
- *Ka' du anbefale dem her?*
- *Ka' I li' muslinger?*
- *Tjah! Det er temmelig overvældende, det her... ikke?*
- *Har du ikke skrevet op, hvad du ska' ha'?*
- *Næ... burde jeg det? [...] Forresten... tak for invitationen!*
- *I kommer vel? BLAP 19-20)*

Talesprogsnært skriftsprog

Efter dette lidt kortfattede signalement af talesproget vender vi nu til-

bage til Højlunds Forsamlingshus og de rigtige talesprogstekster.

Vi skal lave et eksperiment med den tekst, vi begyndte med. Vi skriver den om til skriftsprog, således at alle de mest karakteristiske træk for talt sprog er ændret eller udeladt. I stedet for pauseskråstregerne er der indsat normale kommaer og punktummer. De små signaler til samtalepartneren er udeladt, for i skrevet sprog er det altid forfatteren, der har ordet. Selvrettelserne er fortolket i retning af den mest sandsynlige mening.

Repræsentant for Industrirådet: *Det, der er sket frem til i dag, er rent faktisk, at den U-landsbistand, vi har givet de sidste 20 år, har været med til at kreere nogle hvide elefanter mange steder henne. Men det er ikke det, det handler om. Det, det handler om, det er at komme til at kreere et aktivt erhvervsliv i U-landene.*

Så er der jo nogle, der siger: jamen, har I i industrien ikke også været med til at lave nogle af de der hvide elefanter? Og det har vi måske nok; men vi har spillet med på nogle betingelser, som der er andre, der har sat. Og jeg føler i og for sig, at vi er ved at komme til at leve i en pragmatisk verden, hvor vores interesser i erhvervslivet og U-landsfolkenes interesser kan mødes.

Højlund: *Hvad mente du med 'hvide elefanter'? Var det projekter, der ikke blev til noget, og lignende?*

Repræsentant for Industrirådet: *Ja.*

I denne let bearbejdede version er der ikke længere egentlige fejl set fra et skriftsynspunkt. Men en del talesprogstræk er bevaret, f.eks. sætningskløvnin-
gerne og nogle af fyldeordene. I denne

version er det måske tydeligere end i den originale, hvor megen tomgangssnak ytringen rummer: *frem til i dag, rent faktisk, det det handler om* osv. Det er ord og vendinger uden informationsværdi i sammenhængen.

Det peger på et centralt træk ved samtaleformen i udsendelsen, når man tager i betragtning, at netop denne deltager i Højlands Forsamlingshus havde den længste replik af alle – det ovenstående er kun et uddrag: Man må tale eller blive afbrudt. Det er tydeligt, at der er langt mere respekt for udfyldte pauser af typen *rent faktisk* end for *øh* og *æh* og for udfyldte pauser, som kræver, at den lyttende er opmærksom på finere signaler som f.eks. tonegang. Derfor får en øvet debattør med mulighed for at benytte et politiker- eller embedsmandsagtigt blandingssprog relativt megen tale-tid. Og på dette punkt skal man således ikke anse eksemplet ovenfor for typisk tale.

En mere skriftsprogsnær version

En repræsentant for Industrirådet fremhævede, at dansk U-landsbistand igennem de seneste 20 år har støttet mange 'hvide elefanter', dvs. projekter, der ikke førte til noget eller slet ikke blev gennemført. Han fremhævede det uheldige heri, idet målet er at kreere et aktivt erhvervsliv i U-landene.

Kritikere har påpeget, at industrien også har været med til at støtte sådanne 'hvide elefanter'. Men industriens støtte har været ydet under betingelser, som ikke har været sat af industrien selv. Industriens repræsentant fornemmede dog en stadig større forståelse for pragmatiske synspunkter, således at erhvervslivet og hjælpeorganisationerne lettere kan mødes.

Informationstæthed

Denne sidste version er nærmere typisk skrevet sprog end de to foregående. En af konsekvenserne af dette er, at teksten er blevet noget kortere, selv om den indeholder den samme information. De første versioner er på henholdsvis 179 og 156 ord, denne sidste indeholder kun 96 ord. Det indebærer, at denne sidste version er mere informationstæt end de talesprogsnære versioner.

En konkret beregning af teksternes informationsmængde støtter denne iagttagelse.

Nu er mængden af information i en ytring jo afhængig af modtageren. Hvis modtageren ved det hele i forvejen, så er der ingen information for ham eller hende i teksten; hvorimod der godt kan være mange nye oplysninger for en anden modtager.

En almindelig måde at klare dette problem på er at regne med de enkelte ords hyppighed i tekster. Ideen er, at sjældne ord vil være sjældne for en gennemsnitsmodtager, dvs. at de er mindre forventelige, og at de derfor inderholder mere information end hyppige ord.

Oplysninger om ordenes hyppighed kan man hente i de hyppighedslistes, som tidligere har været omtalt i Mål & Mæle, 8. årg. nr. 1.

Hvis man benytter en optælling af ord i danske fagblade som grundlag, kommer man frem til, at ordene i den første version af debatindlægget gennemsnitligt forekommer 1904 gange i 250.000 løbende ord. Det svarer til en sandsynlighed for, at hvert ord på i gennemsnit 0,76%. I sammenligning hermed er ordene i den skriftnære version væsentlig sjældnere, idet de kun forekommer 1301 gange i gennemsnit af 250.000 løbende ord, svarende til 0,52%. De enkelte ord indeholder således mere information i

den skriftnære version, og den er dermed mere informationstæt.

Også et andet forhold bidrager til at gøre den skriftnære version mere informationstæt end den talte. I den skriftnære version forekommer det enkelte ord kun 1,4 gange i gennemsnit. Derimod gentages ordene meget hyppigere i den talte version, nemlig 2,5 gange. Det er klart, at en lang tekst lettere kommer til at indeholde de samme ord mange gange end en kort tekst. Så det er en overdrivelse at sige, at den talte version gentager sig selv næsten dobbelt så meget som den skrevne. Men der er en tydelig forskel.

Den skrevne, men talesprogsnære version befinder sig *ikke* et sted midt imellem den talte og den skrevne version, hvad angår informationstæthed. Den ligger faktisk helt på linje med den talte version. Den benytter lige så hyppige ord, og de gentages praktisk taget lige så ofte som i den talte version.

Dette betyder blandt andet, at det udmærket kan lade sig gøre at skrive både informationstyndt og informationstæt. Det er en mulighed, det skrevne sprog har, i modsætning til det talte. For hvis man taler ligeså informationstæt, som man kan skrive, kan det være svært for tilhøreren at følge med.

Godt talesprog

Vi har nu set, at talesprog adskiller sig fra skriftsprog på en lang række områder, som alle har at gøre med de vilkår, talen frembringes under.

Det betyder for det første, at *korrekt talesprog* ikke er det samme som korrekt skriftsprog. Man kan ikke sige, at en talt ytring er ukorrekt, blot fordi den ville være ukorrekt på skrift. For det andet betyder det, at *godt talesprog* heller ikke altid er det samme som godt skriftsprog.

Det kan man jo se på alle eksemplerne i denne artikel!

[Forkortelser:

BLAP Panduro: Bertram og Lisa & Anne og Paul (1974)

BM Soya: Brudstykker af et mønster (1946)

DB Riffbjerg: De beskedne III (1976)

DR Danmarks Radio

FV Soya: Frit valg (1948)

HF Højlunds Forsamlingshus

W Riffbjerg: Weekend (1962)

Citaterne fra BM og FV er ændret til moderne retskrivning, dog således at Soyas personlige stavemåder er bibeholdt].

Carsten Elbro

født 1955

Underviser i læsning

ved Københavns Universitet

Erik Hansen

født 1931

professor i dansk sprog

ved Københavns Universitet

Sociale udtalevariationer i København

Efter 2. verdenskrig har man registreret at dialekterne udtyndes mere og mere. Dette har ført til påstande om at alle danskere snart vil tale ens. Men det er ikke rigtigt. Ganske vist er den dialektale tale langsomt ved at forsvinde, men til gengæld opstår der sociale forskelle, dvs. at vores måde at tale på hænger sammen med vores sociale placering. Dette samspil mellem samfund og sproglige forskelle beskrives af *sociolingvistikken*.

Samtidig med at de sociolektale forskelle bliver mere og mere fremtrædende ser vi også en tendens til at den enkelte person varierer sin (ud)tale mere og mere. I vores dagligdag møder vi mange forskellige mennesker i mange forskellige situationer, og vi tilpasser vores sprog til den situation og de mennesker vi er sammen med. Hvis vi ikke gør det så er der nemlig stor sandsynlighed for at omgivelserne vil reagere negativt, eller at dem vi taler med, vil koncentrere deres opmærksomhed om hvordan vi taler (udtryk), i stedet for om hvad vi siger (indhold).



"Vi tilpasser vores sprog til den situation og de mennesker vi er sammen med."

I det følgende vil vi fortælle om (dele af) en fonetisk (lydlig) undersøgelse af sociolektale forhold i København som er blevet foretaget af Projekt Bysociolingvistik, et netop afsluttet forskningsrådsprojekt. Foruden denne artikels forfattere var Jon Albris og Ole Nedergaard Thomsen ansat under ledelse af lektorerne Frans Gregersen og Inge Lise Pedersen. Men først vil vi skitsere Københavns og københavnsks opståen og udvikling.

København og københavnsk

Helt fra 1600-tallet og frem til i dag har København været præget af en massiv indvandring fra landet. Alle landets dialekter kunne dengang høres i byens gadeliv. Vi ved ikke hvad denne sprogforvirring betød for den enkelte indvandrers sprog, om vedkommende fastholdt sin dialekt eller ændrede den. For i forhold til hvilken talesprogsnorm skulle denne ændring finde sted? For at en sprogbrug kan blive til en norm som andre vil efterligne, skal et rimeligt stort antal – måske et flertal – af de omgivende sprogbrugere anvende den. Men i hele det 17. og 18. århundrede var ikke engang 30% af byens indbyggere født i København. Omkring 1850 derimod udgjorde den indfødte befolkning 60% af det samlede indbyggertal. Dette antal er så stort at vi kan regne med, at det er i perioden

umiddelbart op til 1850 at et egentlig københavnsk bliver til.

Da voldene falder bliver de store brokvarterer til, og de befolkes næsten udelukkende med familier fra arbejderklassen. Det er vigtigt at bemærke at ikke alene boede folk der, men de arbejdede der også. Hele deres liv blev levet inden for et meget lille område og med et meget tæt netværk af familie, venner og arbejdskammerater. Det gav og giver mulighed for kvarterspecifikke sprog. Vi ved at indfødte københavnere i 1930'erne var i stand til at høre, om folk kom fra Vesterbro eller Nørrebro.

Middel- og overklassen fjerner sig i samme periode fra de lavere sociale lag, socialt, kulturelt og med hensyn til bopæl. De fortsatte med at bo i centrum, hovedsagelig i Frederiksstad, eller flytter nord for byen eller ud på Frederiksberg. Den højsociale variant af københavnsk får således i denne periode mulighed for at udfolde sig uden at være i tæt forbindelse med den lavsociale variant på broerne.

I slutningen af sidste århundrede oplever København igen en kæmpeindvandring, men på det tidspunkt må vi formode at der var et markant københavnsk som indvandrerne langsomt til lagde sig, når eller hvis de aflagde deres dialekt, og som i hvert fald deres børn fik som talesprog.

Igennem vores århundrede har vi set at København har bredt sig så byen i dag nærmest udgør et område fra Køge i syd til Helsingør i nord og helt til Roskilde i vest. Folks bopæl og arbejdsplads er blevet adskilt, man pendler mellem bopæl og arbejdsplads enten i egen bil eller med offentlig befordring. Det er ikke et problem at bo i Karlslunde eller Espergårde og have sin arbejdsplads i city. På grund af især

70'ernes store saneringer af de gamle brokvarterer (især Nørrebro) er der sket en voldsom udflytning af arbejderklassen, og centrum af København har nu en stor overvægt af studerende (mange fra provinsen) og pensionister, plus en meget stor indvandrergruppe af ikke-danskere.

De to klart forskellige sociolekter, en højsocial og en lavsocial, som fandtes omkring århundredskiftet, er langsomt gået i opløsning i takt med bl.a. den sociale udligning der har fundet sted fra århundredskiftet og frem til i dag, og i takt med den sociale mobilitet. Men hvor gennemført denne opløsning har været, har ikke tidligere været undersøgt, ligesom man heller ikke har vidst om evt. nye sociolektale forskelle er kommet til i perioden.

Undersøgelsens mål og metoder

I 1986 blev Projekt Bysociolingsvistik startet – aldrig har et projekt vist haft et så umuligt navn! Projektets navn skulle afspejle at man forkuserede på bysprog, dvs. københavnsk (fra bydelen Nyboder), og sociolingvistik; det var de sociale forskelle i folks tale der først og fremmest var af interesse. Men ikke nok med det, vi ønskede også at undersøge hvor forskelligt folk taler i forskellige situationer, dvs. hvor stort et stil-repertoire en taler besidder. Vi sagde før at folk til en vis grad skifter tale efter hvem de er sammen med, og efter hvilken situation de deltager i. Det er disse skift vi kalder stilskeft. Det var ikke nok for os blot at få noget tale fra folk; vi ville have fat på den stil der bliver brugt, når den talende overhovedet ikke tænker på hvordan han eller hun taler. Dette vil selvfølgelig aldrig lykkes helt, men man kan forsøge.

Vi har gennemført lange interview, i gennemsnit næsten 2 timer, og vi har båndoptaget folk på deres arbejdsplads og sammen med venner. Interviewerens optræden over for informanten – dvs. personen der bliver interviewet – har på alle måder skullet signalere afslappethed. Således er de fleste af vores optagelser ret formelle i begyndelsen, men efter et stykke tid falder informanten (og interviewer) til ro, atmosfæren bliver mere afslappet, og talen bliver langsomt mere uformel og dagligdags.

De personer vi valgte at bruge i undersøgelsen, skulle være født og opvokset i Nyboder. Fødestedskriteriet sikrede at vi fik det ægte nybodersprog ind på vores kassettebånd, det sprog som en tilflytter til området vil høre og skal forholde sig til. Og vi accepterede kun folk der havde tilbragt hele deres barndom i Nyboder indtil det tidspunkt man "naturligt" flytter hjemmefra, altså omkring 16-18 års alderen.

Før vi gik i "marken" for at udføre feltarbejdet havde vi aftalt hvilke sociale grupper der skulle repræsenteres, og antallet af personer der skulle findes inden for hver gruppe. Projektets opbygning ser ud som følger:

		15-24 årige	25-40 årige
Arbejder- klasse	Kvinder	5	5
	Mænd	5	5
Middel- klasse	Kvinder	5	5
	Mænd	5	5

Vi forsøgte at finde 5 personer i hver af de 8 celler man får, når man todeler de tre sociale dimensioner alder, køn og socialklasse. Målet blev ikke helt nået, men i alle grupper er der mindst 4 repræsentanter. Interview med 31 personer danner grundlag for den fonetiske analyse. Alle analyseresultater bygger på statistiske tests af materialet.

Grænsen mellem 24 og 25 år er ikke så vilkårlig som den måske kan synes. Omkring 25 års alderen har langt de fleste fået et liv der indebærer at de dagligt indgår i mange forskellige situationer. Dette har ikke blot betydning for hvad der bliver sagt, men så sandelig også for hvordan det bliver sagt. Vi er gået et skridt videre: selv i samme sociale situation har vi ikke ment at udtalen var ensartet. Det har vist sig at holde stik. Man kan altså forvente at de 14-24 årige, som alle andre, er fonetisk varierende, men at variationen hos de 25-40 årige måske er større pga. et mere varieret socialt liv.

Denne fokusering på sproglig variation og forsøget på at finde mønstre heri, adskiller vores undersøgelse fra tidligere sproglige undersøgelser af københavnsk.

Vores projekt havde som udgangspunkt en antagelse som sikkert vil overraske mange, nemlig at man i dag i Storkøbenhavn generelt og dermed også i projektets undersøgelsesområde, Nyboder, (stort set) ikke ville finde sociale forskelle i udtalen af vokaler og konsonanter. Denne hypotese kaldte vi vores nulhypotese. Med sociale forskelle tænkte vi først og fremmest på forskelle i udtalen mellem mennesker der er højere og lavere placeret i det sociale hierarki. Vi mente naturligvis ikke at der i dag ikke findes sociolekter i hovedstadsområdet. Som de fleste

kan skrive under på er det som oftest muligt at skelne fx en person fra middelklassen fra en person fra arbejderklassen ud fra måden de taler på. Men der kunne være sket en udligning således at sociolektale forskelle nu om stunder findes på andre niveauer i sproget end det rent lydlige.

Resultater: Gradsforskelle

På baggrund af de fonetiske analyser vi indtil i dag har foretaget af vores nybodemateriale, kan det siges at vores nulhypotese kun delvis blev bekræftet: der er stadigvæk i flere henseender forskelle mellem forskellige socialgruppers udtale af københavnsk, men billedet ikke er så klart som tidligere.

Vi har undersøgt en række af de lydlige træk der er kendt for at markere sociale forskelle. Det har imidlertid vist sig at resultaterne ikke entydigt peger på to klart adskilte sociolekter. Der er i mange henseender sket en opblødning af de gamle grænser, og det bevirker at sprogbrugen ikke længere er forudsigelig i den forstand, at den ene gruppe af personer altid har én bestemt udtale, den anden en lidt anderledes. I dagens København er det mere reglen end undtagelsen at sociolektale særtræk optræder af og til, frem for altid eller aldrig at være til stede.

I og for sig er der ikke noget overraskende i en mere flydende sproglig tilstand, som den her skitserede. Den svarer meget godt til den generelle udvikling byen København har været igennem, og til ændringerne i de sociale forhold i Danmark som helhed. Men den betyder at det er blevet sværere at sige noget generelt om sprogbrugen. Helt overordnet kan det siges at de sociolektale forskelle især viser sig i udtalen af vokalerne. Der findes

også variation inden for konsonanterne, fx i udtalen af konsonanten *t* når den står i begyndelsen af et ord eller en stavelse (med en variation fra et "tørt" *t* til en udgave der falder sammen med et almindeligt *s*), men det er så klart forskelle i vokaludtalerne der dominerer billedet, og det er nogle af dem vi har analyseret og beskrevet.

"Åbning" af *a*, *æ* og *u* i *r*-forbindelser

Vi har undersøgt sproglyden *a* i flere forskellige lydlige forbindelser, jf. udtalen af ordene *kan*, *kaj* og *kam* (se senere om "det flade *a*"). Den opmærksomme lytter vil vide at i ord der skrives med *r*, kan udtalen af vokaler være påvirket af *r*-lyden. Sammenlign fx *ask* med *rask* og *kan* med *kar*. Vi har ikke ment at man kan finde forskelle af sociofonetisk interesse i *r*'ets påvirkning af udtalen af *a*, men vi har undersøgt bl.a. *æ*-lyden (der skrives *e* eller *æ*) før og efter *r*. Tænk blot på en udtale af ordet *ret* der kan lyde som *rat*, og prøv at sammenligne udtalen af *bære* med *bager*. Åbningen af vokalerne *a* og *æ* i *r*-forbindelser er ikke af nyere dato, men den samme bevægelse gør sig nu også gældende ved udtalen af *u* efter et *r*. Fx bliver det mere og mere almindeligt at udtale *russisk* som "rossisk", det samme gælder udtalen "rode" af ordet *rude*. Åbningen *u* til *o* i stillingen efter *r* synes at være opstået i middelklassen, men i dag kan den høres hos alle personer i den yngre aldersgruppe.

"Lukning" af *æ* og *å*

Modsat denne såkaldte åbning af visse vokaler finder vi en lukning i udtalen af andre. Fra starten af vores projekt var vi opmærksomme på at ord som *tænke* og *fængsel* kan høres sagt med bade *æ*

og e, og vi kan af vores båndmateriale se at e-udtalen er i fremmarch, men det er kommet som en overraskelse for os at ændringen æ til e ikke blot kan høres i ord hvor vokalen står foran -ng og -nk i skriften, men også i stillingen foran de to andre nasalkonsonanter -m og -n. Fx kan et ord som *tænde* høres udtalt enslydende med *tinde*, og på lignende måde kan udtalen af *gemme* rime på *skimme*. Variationen synes således nu at være ved at sprede sig til at gælde generelt foran nasalkonsonant, altså ikke alene foran skriftens -ng og -nk men også foran -n og -m. Vi tør dog ikke sige noget sikkert da vores analyse af typen endnu ikke er afsluttet.

Samme forebehold må vi tage over for en anden ny vokaludtale som aflytningen af vores båndmateriale har bragt for dagen. I ord som *også* og *godt* har man længe kunnet høre en spændvidde i udtalen af å-lyden fra en variant der kan lyde i retningen af et a ("Det er gadt"), til en åben og tilbagetrukket å-kvalitet ("Det synes jeg også"). Men det er kommet helt bag på os at flere af vores yngste informanter i den slags ord af og til udtaler en å-lyd der ikke helt falder sammen med det korte å, som man kan høre i *kun* og *ost*, men som går i denne retning.

Opdagelsen af de nævnte nye lydudviklinger i sproget er naturligvis i sig selv spændende, men en mere sammenfattende sociofonetisk beskrivelse kan vi ikke levere endnu.

Derimod tør vi godt fremlægge følgende analyse:

Det flade a

I Københavns Radio kan man høre et ugentligt tilbagevendende program blive annonceret "Med i weekenden". Den gastronomisk interesserede ved at

der menes *mad*, men *med* kunne også give god mening – som overskrift til et andet program. Det der er interessant i denne forbindelse, er om man ud fra udtalen af ordet *mad* med fladt a kan slutte noget om speakerens sociale gruppetilhørsforhold, fx om det er en person fra arbejder- eller middelklassen der taler. Svaret er nej, har vores undersøgelser vist. Man kan ikke skelne én gruppe af personer fra en anden på basis af deres udtale af det lyse a.

Man finder lyden i ord med en bestemt opbygning, nemlig – set gennem retskrivningen – dem der ender på skrevet a, fx i *da*, og i ord hvor a'et står foran ét af konsonantbogstaverne t, d, s, l og n, altså fx i *hat* og *han*.

For det første er der ingen af vores informanter der udelukkende bruger én og kun én variant af det lyse a, og hermed er det også sagt at vi ikke er stødt på en person der kun anvender det flade a. Faktisk anvender alle informanterne de to a-varianter som skrivemåderne *med* og *mad* antyder, nogle kan også finde på at udtale det lyse a en anelse længere tilbage i munden, dvs. i retningen af den a-nuance man hyppigst hører i et ord som *vej*. Denne variation mellem forskellige a-nuancer finder man vel at mærke ikke blot i afslappet tale med kammerater eller arbejdskolleger, men også i uafslappet eller formel talestil.

For det andet høres de forskellige a-lyde omtrent lige hyppigt (eller sjældent) hos yngre og ældre (husk vores "ældre" er 25-40 år gamle), hos kvinder og mænd, og i arbejderklassen og i middelklassen. Den ene gruppe har ikke noget at lade den anden høre.

Og endelig, for det tredje, er det flade a i intet tilfælde den a-variant der optræder tiest. Alle kender det flade a

som en udtalemulighed, alle siger den af og til, men den dominerer ikke i hyppighed over det mere åbent udtalte a.

Sammenligner vi samme personers udtale i de to talestilarter vil har undersøgt: afslappet og formel stil, så viser det sig for alle grupper, og inden for hver gruppe også for flertallet af enkeltindivider, at det flade a er sjældnere i formel tale sammenlignet med uformel talestil, men i samme talestil – og det er pointen hér – er der ingen forskel på folk mht. til den relative forekomst af varianter af det lyse a.

Dette resultat skal nok overraske den der er af den faste overbevisning at arbejderklassen er hovedrepræsentanten for det flade a. Arbejderklassen var først ude med denne a-kvalitet, men den er altså i dag (i den voksne generation, jf. nedenfor) blevet fællesje.

Mod større forskelle

Når vi ikke generaliserer den sociale enighed i brugen af det lyse a 100% så skyldes det at vi har fundet tegn på, at dele af den helt unge generation er ved at genindføre en tilstand der peger tilbage mod den tidligere herskende forskel mellem socialklasserne mht. brugen af de korte, lyse a-nuancer: Den opvoksende middelklasse er, i modsætning til deres arbejderklassekammerater, mindre tilbøjelig til at anvende det flade a end forældregenerationen er, og det er ganske interessant at det kan se ud til at det er de unge mænd i gruppen der fører an i denne modtendens. Vi tør godt vove den forklaring at vi hér har et nykonservativt træk i sproglig klædedragt. Mens 60'erne generationen synes at dele brugen af det flade a imellem sig (man kan måske sige at de to socialklasser i denne hen-

seende kommer hinanden i møde), er der altså nu tegn på at børnene af denne generations middelklasse – muligvis anført af de unge mænd – er på vej mod at lægge en dæmper på og at modarbejde denne sociale lighed i brugen af det flade a.

Projektet Bysociolingvistik's nulhypotese gælder altså som en sociofonetisk beskrivelse af sprogtilstanden for det korte, lyse a i hovedstadsområdet i 1980'erne, men billedet vil muligvis ændre sig i fremtiden. Måske er den nys forsvundne sociolektale opsplitning mellem arbejder- og middelklassebrugere af denne lydtype på vej tilbage, så brugen af den igen vil få social betydning.

Stilvariation

Sammenligner vi hver enkelt persons tale i hver af de to stilarter, afslappet og formel tale, så finder vi forskelle. Tager vi alle vores informanter under ét, er de enige om at bruge færre flade a'er når de skal tale pænt. Ser vi nærmere på hver enkelt af de 8 grupper som vores informanter maksimalt kan inddeles i, viser det sig at det – når man ser bort fra de 14-24 årige middelklasse-mænd, der som nævnt ovenfor kan betragtes som en gruppe i bevægelse – er den ældre aldersgruppe (de 25-40 årige) der bærer stilskiftet. Denne kendsgerning synes at bekræfte at stilvariation almindeligvis først indarbejdes når et mere varieret socialt liv kræver det. Blandt de 25-40 årige er det ganske klart at det er kvinderne, i særdeleshed dem fra middelklassen, men også dem fra arbejderklassen der er mest tilbøjelige til at droppe det flade a i mere formelle situationer, og lige efter kvinderne kommer "gruppen i bevægelse", de unge middelklasse-mænd.

Det er almindeligt kendt i den sociolinguistiske litteratur at kvinder normalt er mere følsomme over for sociale signaler i sprog end mænd er, og at kvinder i deres brug og vurdering af sprog plejer at placere sig på den "pæne" side (= den konservative). Der er således ikke noget påfaldende i at det er kvinderne der har størst sproglig variation fra ét stilniveau til et andet. Stilvariationen hos de unge middelklassemænd er derimod isoleret betragtet opsigtsvækkende, men sammenholdt med vores analyse af forskellene mellem socialgrupperne bekræfter den vores fortolkning af denne gruppe som spydspidsen i en genindførelse af sociale forskelle i brugen af det korte, lyse a.



Sammenfatning om det lyse a

Sammenfattende om det korte, lyse a kan det siges:

- 1) at der stort set ikke er sociale forskelle når stilen holdes konstant,
- 2) at a-varianterne alligevel for medlemmerne af sprogsamfundet har (bibeholdt) en social betydning der kommer frem til overfladen, når man sammenligner forskellige stilniveauer,

3) at der synes at være tegn på at en sociolektal tvedeling i arbejder- og middelklassebrugere igen vil opstå, og

4) at der ikke er antydninger af at det flade a vil blive den dominerende a-variant.

Gå ud og lyt

Når den sociale virkelighed er så uigenkendskuelig som den er i dag, bliver de sproglige analyser tilsvarende komplicerede når man ønsker at sammenholde udtalen med virkeligheden. Dette bør naturligvis ikke forhindre nogen i selv at lytte nøje til den tale der omgiver dem, men man skal blot ikke længere regne med uden videre at kunne slutte fra et enkelt fladt a til den talendes sociale identitet.

Henrik Holmberg, f. 1945
Forskningsstipendiat
Københavns Universitet

Erik Møller, f. 1958
Kandidatstipendiat
Københavns Universitet

Tekst, tale og højtlesende maskiner

Som bekendt foregår menneskers kommunikation med datamater almindeligvis med skrevet tekst som bindeled. Vil man have en datamat til at udføre et arbejde – fx at lade et tekstbehandlingsprogram fra lageret ind i arbejdsbureauet, indtaster man en bestemt tekst, typisk programmets navn, på det tastatur der er forbundet med datamaten, hvorefter det program der altid fungerer som baggrundsprogram i en tændt datamat – kommandofortolkeren i datamatens operativsystem – sørger for det videre fornødne. Når tekstbehandlingsprogrammet så er sat i gang, er det datamatens tur til at meddele brugeren, at den er klar til at modtage nye ordrer, og det sker ligeledes ved hjælp af skrevet tekst – eller i det mindste visuelle signaler – denne gang på en skærm.

Denne kommunikationsform har fordele og ulemper. I visse situationer er det naturligvis en fordel at datamaten meddeler sig ved skrevet tekst eller ved at tegne på en skærm, men det er på sin vis et primitivt træk ved datamater at skrevet tekst er deres *eneste* kommunikationsform, og at det oven i købet er et yderst begrænset udvalg af konkrete, korte tekster datamatens programmer kan give fra sig på skærmen, eller fortolke når de taster ind. Menneskers naturligste og almindeligste kommunikationsform er det talte sprog, og det har mange fordele. For blot at nævne to af de mest indlysende: man kan tale 'om hjørner', dvs. man behøver ikke at være inden for fysisk

eller visuel rækkevidde for den man taler til eller den der taler til én, og man skal ikke gå rundt med klodsede apparater som tastaturer og skærme for at kunne afgive og modtage meddelelser.

Højtlesning

Det er derfor ikke underligt at man i mange år har forsket i hvordan man kan frigøre kommunikationen med datamater fra bindingen til skrevet tekst. Idealet er at brugeren skal kunne tale til en datamat, og at datamaten skal kunne tale til brugeren. Der er ikke udset til at dette ideal vil blive nået inden for en overskuelig fremtid, men det er i dag muligt til en vis grad at få datamater til at simulere *højtleseprocessen* , nemlig på den måde at den skrevne tekst der normalt sendes til en skærm, bliver omsat til talesignaler der sendes til en højtaler. Der er med andre ord tale om en proces der i et vist omfang er udstyret med sproglig 'viden'.

Det at læse en sætning højt er en kompliceret proces, som formentlig i det mindste omfatter følgende elementer.

- 1) *Genkendelse* af tekstens ord.
- 2) *Sproglig fortolkning* af de genkendte ord og deres arrangement.
- 3) *Planlægning* af udtalen af den genkendte og fortolkede tekst.
- 4) *Udførelsen* af den lagte plan i form af 'ordrer' til taleorganerne og disses 'efterkommelse' af ordrene.

Hvis man antager at de nævnte elementer i en eller anden forstand er til stede, er det klart at der er mange frihedsgrader i den samlede proces, og det er indlysende at jo flere frihedsgrader der er, desto flere muligheder er der for hvordan resultatet – den oplæste sætning – kommer til at lyde.

Ordgenkendelse

For genkendelseselementets vedkommende ligger frihedsgraderne især i at skrevne ord – anskuet som bogstavfølger uden mellemrum – i mange tilfælde er flertydige, hvis de betragtes isoleret. Læsere af *Mål og Måle* husker måske Carsten Elbros artikel *Får fortænder fortællinger* i nr. 1, 1987 p. 13-15, hvor denne egenskab ved skrevne ord eksemplificeres med ord som *skovrider*, *blegrød*, *føderation*, og naturligvis den klassiske *trækvinde*.

Det er et interessant træk ved læsning i almindelighed, at sådanne ord normalt overhovedet ikke volder problemer, og at læsere typisk slet ikke bemærker flertydigheden, men straks vælger en bestemt – den 'rigtige' – genkendelsesmulighed. Hvis genkendelsesproblemet spiller en rolle i et automatisk højt-læsesystem, kan man derfor sige at systemet på dette punkt simulerer den menneskelige læseproces dårligt.

Hvad angår fortolkningselementet, ligger frihedsgraderne især i grammatiske og betydningsmæssige flertydigheder, der hidrører fra at skrevet tekst er ret fattig på sproglig information af mere subtil art. En skrevet sætning som *Jensen stod på bussen* svarer til flere forskellige talte sætninger. Hvis betydningen er a): 'Jensen steg ind i bussen' udtales udsagnsordet tryk-

svagt; hvis betydningen er b): 'Jensen befandt sig stående (oven) på bussen', udtales det med et vist tryk, der indgår i sætningens trykmønster på helt samme måde som trykket på *trak* i *Jensen trak i snoren* indgår i denne sætningens trykmønster. Den forskel det her drejer sig om, vil mange sprogfolk tilskrive en underliggende forskel i sætningernes grammatiske mønster, og trykforskellen i udtalen vil ud fra denne betragtning blive beskrevet som en regelmæssig *konsekvens* af denne mønsterforskel. Selve den skrevne sætning har imidlertid også i det mindste tre andre fortolkninger og dermed forbundne udtaler, hvor det er naturligere at sige, at det *grammatiske* mønster er som a) eller b), men det *betydningsmæssige* mønster er forskelligt. Man kan lægge stærkere tryk på *stod* og derved modificere betydningsvægtingen i forhold til b) i retning af c): 'Det var stående – ikke fx liggende eller siddende – Jensen befandt sig (oven) på bussen'; eller man kan lægge sætningens stærkeste tryk på forholdsordet *på* og dermed modificere vægtingen af betydningen i forhold til a) i retningen af d): 'Jensen steg ind i bussen, ikke fx ud af den', og i forhold til b) i retning af e): 'Det var oven på bussen Jensen stod, ikke fx ved siden af den eller inde i den'. Trykforhold som de sidstnævnte, der er eksempler på det man kan kalde *værditryk*, følger faste mønstre, der er lige så regulære bestanddele af talt dansk som fx ordstillingsregler. Men de er normalt ubetegnede i skrift, og i situationer hvor skrevet tekst er 'input', er det derfor overladt til oplæseren - hvadenten det er et menneske eller en maskine – at vælge fortolkning. Men når en fortolkning er valgt, er visse dele af udtalen, altså fx trykfordelingen, dermed bestemt.



Jensen stod på bussen – I og II

Omformning

Genkendelse og fortolkning udgør tilsammen det man kan kalde *bestemmelse* af hvilket sprogligt mønster man har for sig i teksten, mens planlægning og udførelse er *omformninger* af givne mønstre til talesignaler. Også i disse elementer er der frihedsgrader. Planlægningsfasen svarer til valget mellem flere mere eller mindre ligeberettigede måder at udtale et sprogligt mønster på. I alle de nævnte fortolkninger af *Jensen stod på bussen* må man som oplæser vælge om man vil udtale det bløde *d* i *stod* eller udelade det. De to udtaler af dette ord er ligeberettigede i den forstand at sætningen betyder nøjagtig det samme uanset hvilken udtale man vælger, selv om samme person

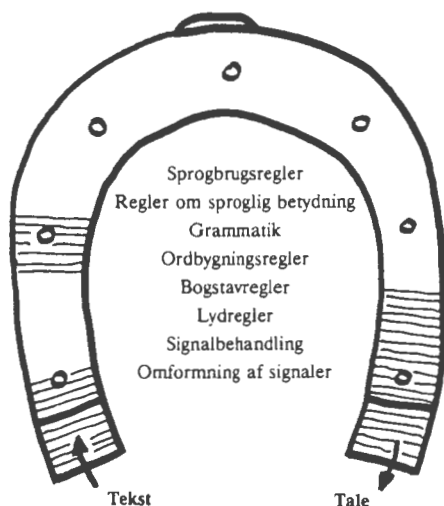
måske vil være tilbøjelig til altid at vælge den ene.

Inden for udførelsen er frihedsgraderne også mange, men de er her mindre følelige, fordi de især angår 'mekanisk slør', dvs. små afvigelser i detailindstillinger af taleorganerne fra gang til gang samme person siger samme ytring.

De omtalte elementer må naturligvis ikke forstås i alt for konkret betydning, og de må frem for alt ikke opfattes som fire forskellige faser i højt-læsningsprocessen. De skal opfattes som forskellige sider af processen, dele som griber ind i hinanden og giver hver deres bidrag til såvel bestemmelse af mønstret som omformning til tale. En god model for automatisk (datamatisk) omformning af skrevet tekst til tale må tage hensyn til alle de nævnte elementer og til samspillet mellem dem. Jeg vil i det følgende præsentere en generel model for datamatisk bestemmelse og omformning af sprogligt input. Denne model har nogle rammer der åbner mulighed for at indbygge programmer til simulering af hvert af de omtalte elementer (og flere til). Efter præsentationen af denne model vil jeg omtale det indhold modellen har eller bør have, når inputtet er skrevet tekst og outputtet syntetisk tale, dvs. 'tale' frembragt uden at en menneskestemme har været indblandet i processen.

En model for datamatisk sprogbehandling

Modellen kan anskueliggøres som en hesteskoformet figur som den viste, hvor det sproglige input kommer ind for neden til venstre, går rundt i hesteskoen og kommer ud i et andet format nede til højre.



Venstre del af hesteskoen svarer til *mønsterbestemmelsen* af inputtet, og højre del til *omformningen* til det ønskede output. De forskellige navngivne lag inde i hesteskoen viser de forskellige sproglige og tekniske specialområder datastrømmen kommer i berøring med på vej rundt. Hvad der sker i de enkelte områder afhænger dels af input og output, dels naturligvis af hvor meget sproglig og anden 'viden' der lægges ind i modellen. I det konkrete danske højt læsesystem jeg især vil basere min gennemgang på, foregår der mest i områderne *regler om ordbygning* og *grammatik* i venstre side, dvs. mønsterbestemmelsesdelen. I højre side, omformningsdelen, foregår der mest i områderne *bogstavregler*, *lydregler* og *signalbehandling*. Signalbehandling vil i tekst-til-tale systemer sige dannelsen af elektroniske mønstre, der svarer til lydbølger. Disse områder er markeret ved skravering. Det er disse områder og de opgaver de skal løse jeg vil omtale i det følgende. Jeg vil dog lægge hovedvægten på de sproglige komponenter i venstre side af hesteskoen og kun komme meget

kort ind på komponenterne i højre side, ikke fordi de ikke er vigtige, men fordi det alligevel ikke er muligt her at behandle disse temmelig tekniske komponenter på fyldestgørende måde. Desuden er det i mønsterbestemmelsesdelen de største *sproglige* udfordringer ligger, selv om de *lydlige* regler og signaldannelsen naturligvis er af stor betydning for det generelle *lydlige* indtryk af den automatisk dannede tale.

I begge sider af hesteskoen er der – som skraveringen viser – også en del virksomhed i det område der kaldes *omformning*. Omformning bruges her om den formændring sprogligt materiale må underkastes (i inputhjørnet) for overhovedet at kunne behandles af en datamat, nemlig omformning til elektroniske mønstre, og den formændring der skal til for at omsætte elektroniske mønstre til 'naturlige' fremtrædelsesformer som fx skrevet tekst eller tale (i outputhjørnet). Omformning af lydlige signaler til elektroniske mønstre som ved talt input kaldes AD-Konvertering (analog-til-digital), og den omvendte proces kaldes DA-konvertering. Omformning af skrevet tekst i form af trykte bogstaver til elektroniske mønstre kan i dag med ret stor pålidelighed foretages af såkaldte tekstscannere. Pendanten til DA-omformning når *outputtet* skal være skrevet tekst, er den proces der foregår i en printer der er koblet til en datamat som ydre enhed.

Mønsterbestemmelse

I områderne for orddannelse og grammatik simuleres ordgenkendelse og fortolkning, og jeg vil først omtale disse elementer i processen. At mennesker kan genkende ord på deres modersmål er en trivial kendsgerning, der

tyder på at det hører med til almindelig sprogbeherskelse at have et ordforråd. Det er derfor oplagt at give en maskine der skal kunne læse højt, et ordforråd, dvs. opbygge en *elektronisk ordbog* ved at lægge en stor mængde ord ind som datamønstre, der rummer oplysninger om ordenes egenskaber. Spørgsmålet er imidlertid hvilken slags 'ord' og hvor mange ord den skal have. Det er velkendt at et menneskes ordforråd ikke er særlig veldefineret, og det er sandsynligt at de ord et menneske har til rådighed, ikke nødvendigvis udgør en ensartet gruppe ud fra et sprogligt synspunkt, altså fx at det udelukkende er rødder eller endelser eller udelukkende hele ord uanset om de er sammensatte. Snarere er det en blanding, således at fx et simpelt ord som *sten* og et sammensat ord som *lagkage* i en vis forstand har samme status i ordforrådet, selv om de fleste godt kan indse at de i en vis anden forstand har forskellig status – nemlig som hhv. et simpelt og et sammensat ord. Når man skal fastlægge det ordforråd man gerne vil tildele et automatisk system, må man nødvendigvis lægge en skillelinie mellem ord der skal med og ord der ikke skal med, og den skillelinie man vælger vil afhænge af hvad det automatiske system skal kunne.

Hvis det drejer sig om en højt læse-maskine kan man med fordel lægge hovedskillelinien mellem netop simple og sammensatte ord. For det første er mængden af simple ord lettere at afgrænse end mængden af 'ord' ud fra andre kriterier. For det andet kan det meste af den information der er nødvendig for at bestemme netop *udtalen* udledes af de egenskaber der er knyttet til de enkelte rødder eller endelser i et konkret tekstord; man vil derfor typisk lægge ordene *lag* og *kage* ind, men ikke

lagkage. For at kunne bestemme det sidste ords udtale må systemet til gengæld være udstyret med et *program* der kan opløse ordet i dets bestanddele og i denne forbindelse opsamle de oplysninger der er knyttet til de to genkendte sammensætningsled, for *lag*'s vedkommende bl.a. den oplysning at ordet har forkortelse af vokalen (ledsaget af mørkfærvning af a'et) og stødtab når det er det første led i en sammensætning. Det er derimod ikke relevant for *udtalen* at ordet *lagkage* har en speciel betydning, der ikke – i hvert fald ikke udelukkende – kan siges at være kombinationen af leddenes betydninger.

Ståltråd

Nu er det ikke i alle tilfælde så simpelt at man kan nøjes med at lægge simple ord ind i ordbogen – og naturligvis den lille faste mængde af bundne forstavelser og endelser. Ordene *ståltråd* og *stålstang* har begge *stål*- som førsteled men forskellige sidsteled. Men i *ståltråd* er der ligesom i *lagkage* forkortelse og ændring af vokalen samt stødtab i førsteledet; i *stålstang* og andre sammensætninger med *stål*- er der kun stødtab. Man kan derfor ikke nøjes med at lægge *stål* ind med den oplysning at det har stødtab som førsteled i en sammensætning, da det (i langt de fleste tilfælde) ville føre til en utilsigtet udtale af *ståltråd*. Man må lægge ordet *ståltråd* ind som helhed med de relevante oplysninger, og desuden indrette programmet til ordanalyse sådan at ordet faktisk bliver genkendt som ét ord, og ikke som sammensat af ordene *stål* og *tråd*. Forholdet er i øvrigt mere indviklet end som så: *Ståltråd* i den udtale jeg her omtaler – dvs. rimende på 'troldkåd' – har en ret snæver betydning, hvori bl.a. indgår, at den ikke be-

høver at være af stål og at dens diameter er 1 - 2 mm.

Ståltråd rimende på '*stålbåd*' ville derimod både udtale- og betydningsmæssigt være en sammensætning på linie med *stålstang*; det ville man fx kunne kalde centimetertyk tråd af stål. Det er derfor ikke i *alle* tilfælde rigtigt at fortolke *ståltråd* som den leksikaliserede enhed.

Af andre genkendelsesproblemer kan nævnes ord af *trækvinde*-typen. Problemets løsning i et automatisk system afhænger af ambitionsniveauet. For de fleste ord af denne type gælder det, at det næsten altid er den samme mulighed der er den rigtige i tekster, og man kan så indrette systemet sådan at det netop vælger denne, for eksempel – som ved *ståltråd* i den almindeligste betydning – ved at lægge den hel ind i ordbogen. Denne løsning er praktisk, og hvis systemet ikke skal inddrage grammatiske og betydningsmæssige kriterier i forbindelse med mønsterbestemmelse, er det formentlig den bedste løsning.

Grammatiske mønstre

Hvis systemet derimod er udstyret med en komponent der repræsenterer grammatisk indsigt, kan man ved at anvende denne indsigt hensigtsmæssigt i en hel del tilfælde træffe et entydigt valg mellem flere muligheder af denne art. Det gælder fx *trækvinde*, der jo med den ene fortolkning er ubestemt form flertal af *trækvind*, med den anden ubestemt form ental af *trækvinde*. Hvis der i den skrevne tekst der skal fortolkes er knyttet et talord, et tillægsord eller et kendeord til *trækvinde* vil dette (i korrekt dansk) ved sin bøjningsform vise om der er tale om det ene eller det andet ord, jf. en skrevet

sætning som *hun er en rigtig trækvinde*. Men med *trækvinde* er man lige vidt, da *træk-vinden* og *træk-kvinde* kan forekomme i de samme grammatisk definerede sammenhænge: man kan ikke ud fra grammatiske kriterier foretage en entydning/genkendelse af dette ord i en skrevet sætning som *det var trækvinde der smækkede vinduet i*.

I det konkrete system for dansk tekst-til-tale er den grammatiske komponent for tiden netop på et udviklingsstrin hvor entydiggørelser på grundlag af overensstemmelse i køn og tal mellem navneord, tillægsord og kendeord kan gennemføres i tilfælde hvor den pågældende information er til stede i teksten. Men det er vigtigt at understrege at den grammatiske analyse der gør dette muligt ikke udelukkende er lagt ind i systemet for at kunne foretage entydige *ordgenkendelser* som de nævnte. Dette ville nemlig ikke være ulejligheden værd til praktiske formål (som fx konstruktion af højtælsemaskiner for blinde), hvis der ikke var andre problemer i mønsterbestemmelsen der i nok så høj grad krævede grammatisk analyse. Korrekt genkendelse af enkelte bøjningsformer af visse sjældent forekommende ord ville bidrage meget lidt til en forbedring af et sådant systems kvalitet, bl.a. fordi forkerte genkendelser i mange sådanne ord kun ville betyde små forskelle i udtalen; der er fx kun to *lydlige forskelle* mellem de to udgaver af *trækvinde*: /trækvenen/ og /trækven?en/ ('?' angiver stød).

Trykmønstre

Der er imidlertid andre problemer ved mønsterbestemmelse på grundlag af skrevet tekst som kræver et vist mål af sætningsanalyse for at kunne løses til-

fredsstillende. Det drejer sig om fastlæggelsen af nogle mønsteregenskaber ved danske sætninger der er afgørende for sætningernes tryk. Det er vigtigt at disse problemer løses, fordi de trykforhold det drejer sig om forekommer i næsten enhver sætning, og derfor i ganske særlig grad bidrager til det lydlige særpræg ved talt dansk.

Et af mine yndlingseksempler på sådanne mønsterbetingede trykforhold er sætningerne

- 1 Jensen jager.
- 2 Jensen jager mus.
- 3 Jensen jager musene.
- 4 Jensen jager musene ud.
- 5 Jensen jager musene ud fra personlige motiver.

Ingen med dansk modersmål (og en smule sprogøre) vil formentlig benægte at udsagnsordet har tryk i sætning 1, 3 og 5, men er tryksvagt i sætning 2 og 4.

Der kan som bekendt gives regler for tryktabet i sætning 2 og 4. Det kan diskuteres – og er blevet diskuteret af flere – hvordan sådanne tryktabsforbindelser skal beskrives betydningsmæssigt, men det er et vigtigt og i mange sætninger entydigt *symptom* på forbindelser af typen i sætning 2, at der er tale om udsagnsord med et genstandsled i nøgen form, dvs. i ubestemt form og uden nogen foranstående indkredsende bestemmer i form af et ubestemt kendeord eller et talord. Sammenlign fx sætning 2 med sætning 6 og 7:

- 6 Jensen jager en mus.
- 7 Jensen jager tre mus.

På samme måde kan tryktabet i sætning 4 (i modsætning til sætning 5) be-

skrives ved at udsagnsordet og biordet *ud* indgår i en forbindelse med enhedstryk.

Pointen er at et automatisk system der råder over en 'ordbog' og dermed har adgang til informationer af typen 'kan tage genstandsled' (*jage*), 'er en indkredsende bestemmer' (*en, nogle, ect.*) kan bringes til at fortolke sådanne sætnings mønster entydigt, hvis det forsynes med regler der udtrykker disse sammenhænge mellem egenskaber ved ordformer og deres rækkefølge i sætninger, altså med hvad nogle ville kalde overfladegrammatiske regler. Den disciplin der går under navnet *datamatisk lingvistik* beskæftiger sig bl. a. med hvordan man kan formalisere sådanne regler, så de kan gøres fortolkelige for et datamatprogram.

Omformning

Uanset hvor langt et givet system når op i hesteskoen på venstre side, kommer der et punkt hvor den information der er resultatet af inputtets beskaffenhed og den genkendelse og fortolkning af det som tilføres fra de forskellige sproglige komponenter, skal udmøntes på højre side som en omformning til det format der ønskes i outputtet, altså i dette tilfælde syntetiske talesignaler.

Som det fremgår af figuren er områderne *bogstavregler*, *lydregler* og *signalbehandling* de vigtigste i denne side af modellen. I et system som det danske, hvor den sproglige information tilføres systemet ved som ovenfor omtalt at sammenholde ordbogsinformationer knyttede til rødder eller endelser med grammatiske regler, er det i praksis mest bekvemt at lagre den for udtaalen relevante information i form af *ord i grov lydskrift*, dvs. som en slags notation ud fra hvilken det enkelte ords lyd-

lige opførsel kan bestemmes entydigt i alle sammenhænge. Et eksempel vil gøre dette klarere: efter mønsterbestemmelsen har sætning 2 og 3 ovenfor følgende fremtrædelsesformer:

2a jenhs ,On jag*Or%mus
3a jenhs ,On jag*Or mus+OnO

Af de symboler der indgår i disse notationer hidrører de fleste fra de genkendte ordbogensenheder. Kun mellemrummene og %-tegnet i 2a hidrører fra den grammatiske programdel. Det man skal bemærke her, er at enhedstrykket i 2a i denne notation signaleres af %-tegnet mellem jag*Or og mus. Det er en opgave for programdelen for bogstav- og lydregler at omforme disse notationer til lyde, dvs. en fin lydskrift, og dette sker ved at de underkastes et sæt lydlige omformningsregler i en bestemt rækkefølge. Effekten af denne lydlige omformning er at de to sætningers notationer bliver ændret til

2b 'jEnsOn ja:jC'mu':s
3b 'jEnsOn 'ja:jC 'mu':sOnO

I den viste form overtages sætningerne af den *lydlige programdel*. Først i denne del af systemet tildeles de enkelte symboler *lydlige* værdier, dvs. nogle talsæt, der står for de lydlige karakteristika ved de lydelementer symbolerne repræsenterer. Herefter underkastes disse rækker af lydlige elementer et nyt sæt omformningsregler, der varetager de forskellige lydlige justeringer og udglatninger der er karakteristiske for sammenhængende tale.

Den fremtrædelsesform de to eksempelsætninger derefter har, er principielt af samme art som i den lydlige programdel, dvs. en række af 'pakker' af talværdier, men der er typisk flere

elementer end de der er repræsenteret i en lydskrift. Disse talrækker bliver nu sendt videre til *signalbehandlingsdelen*, der i dette tilfælde er en syntesimulator, dvs. et program der simulerer effekten af et sæt lydlige filtre på de svingningsmønstre som tallene repræsenterer.

Effekten af dette er et meget stort sæt elektroniske mønstre der repræsenterer et lydligt bølgetog, og via den ovenfor omtalte DA-omformning bliver det endelige resultat et elektrisk svingningsmønster, der kan drive en højttaler.

Det tjener ikke noget formål i denne sammenhæng at gå nærmere ind på de sidstomtalte, mere teknisk betonede dele af omformningsprocessen. Men jeg håber at jeg har kunnet give et indtryk af hvor kompliceret hele denne mildt sagt avancerede form for tekstbehandling er.

Peter Molbæk Hansen f. 1946
Adjunk i Datalogivistik
ved Københavns Universitet

Sprogviden

1. Hvordan udtales det engelske ord *steak*? [stiik] eller [stæjk]?
2. Hvilke af de følgende ord skal staves med æ:
klokkesl#t – h#fte – d#lle – #gypter
3. Hvilke af følgende sætninger skal *udenom* skrives i to ord:
 - a. Læg et stykke papir udenom!
 - b. Behøver vi at gå udenom skoven?
 - c. Dette er en bog, som man ikke kan komme udenom.
 - d. Har du ikke et stykke papir at lægge udenom?
4. I hvilke af de følgende sætninger er der kommafejl:
 - a. Jeg synes, vi skal ringe, når vi kommer hjem.
 - b. Det, han mente, var, at de, der ikke vil med, heller ikke behøver at betale.
 - c. Ham tror jeg ikke vi skal vente at se igen.
5. Udpeg de rigtige stavemåder af følgende tre ord:
billed/billede – hoved/hovede – hundred/hundrede
6. Af følgende engelske ord i dansk har kun ét samme betydning i dansk og i engelsk. Hvilket?
speaker – choker – speeder – smoking – spencer
7. Hvilke to ord af følgende er historisk beslægtede:
gren – gro – grøn – grå
8. Ingen af de følgende sætninger indeholder sprogfejl i egentlig forstand. Men en af dem indeholder noget sprogligt vrøvl. Hvilken?
 - a. Hen på eftermiddagen tog stormen noget af.
 - b. Merudgifterne blev over 2 mio. kr. højere end beregnet.
 - c. Nu gælder det om at få truffet aftaler med alle, som skal deltage i turen.
9. I hvilke af følgende sætninger regnes *som der* for korrekt:
 - a. Han hører til den slags mennesker, som der ikke kan udsættes noget på.
 - b. De kasser, som der er tilbage, skal stilles ud ved gavlen.
 - c. Argumentet er nøjagtigt det samme, som der blev brugt sidste gang.
10. Hvad betyder det græske fremmedord *rhabdomantik*:
 - a. Læren om fortællende digtning.
 - b. Kunsten at finde skjulte vandårer med en pilekvist
 - c. Sindssygdом, som ytrer sig ved ubehersket vildskab.

Løsninger findes s.8