

## Tekstanalyse og kompetanse

Distinksjonen mellom begrepshierarkier (som leksikografer og terminologer er interessert i) og temahierarkier (som anvendes innenfor informasjonsgjenfinningssystemer) er viktig. Hitil har vi omtalt og fokusert på forskjellene mellom begrepshierarkier og temahierarkier. Vi har imidlertid ikke diskutert hvilket temabegrep man kan anvende. Hva er tema i dokumenter? Hvordan kan man isolere, identifisere og analysere tekstuelle temaer?

I det følgende skal vi omtale temabegrepet slik det er behandlet innenfor tekstlingvistikken. Hvis man skal gjøre seg håp om å kunne svare på spørsmålene ovenfor, må man også ta stilling til hva en tekst er. Kan tekster for eksempel sies å være en grammatisk størrelse? Dette spørsmålet er i sin tur avhengig av hva man legger i begrepet grammatikk.

Ut fra dette er det naturlig å ta utgangspunkt i det lingvistiske kompetansebegrep, og i lys av dette prøve å avgrense det studieobjektet som skal undersøkes: analyse av tematiske enheter i dokumenter/tekster.

La oss begynne med følgende spørsmål: Hvor i det lingvistiske terrenget befinner vi oss når vi skal analysere teksttematiske enheter i dokumenter i lys av informasjonsgjenfinning? Dette spørsmålet er det verd å gå nærmere inn på. I det følgende skal vi derfor drøfte kompetansebegrepet i denne sammenhengen, og hvilken relasjon kompetansebegrepet har til vårt temabegrep. Det vil forhåpentlig kunne sette oss i stand til å krysse av på det lingvistiske kartet hvor vi befinner oss. Det kompetansebegrepet vi legger opp til vil også kunne gi oss en beskrivelsesramme for å drøfte begreper som *tekst*, *koherens*, *temaer* og *makrostrukturer*.

I moderne lingvistikk er det (særlig) to kompetansebegreper som diskuteres. Noen betrakter dem som uforenlige, andre som gjensidig utfyllende begreper. Lyons f.eks.

tilhører den sistnevnte kategori. Det ene av disse to begrepene er Chomskys *lingvistiske kompetansebegrep*:

## F

*Linguistic theory is concerned primarily with an ideal speaker - listener in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance (Chomsky 1965:03).*

Det andre kompetansebegrepet studeres ofte under betegnelsen *kommunikativ kompetanse*, og inkluderer normalt Chomskys lingvistiske kompetanse. Betegnelsen stammer opprinnelig fra Hymes (1971), som definerer det som en persons kunnskap om, og evne til, å anvende de semiotiske systemene som er tilgjengelige for ham/henne i et språksamfunn.

Den lingvistiske kompetanse er altså inkludert i den kommunikative. Studiet av den lingvistiske kompetanse kaller Lyons mikrolingvistikk, til forskjell fra den mer omfattende sosiolingvistikken eller stilistikken (av Lyons også kalt makrolingvistikk). En viktig del av den kommunikative kompetanse som faller utenfor den lingvistiske kompetanse er den innfødte språkbrukerens evne til å kunne *kontekstualisere* systemsetninger.

Begge de to typene kompetanse refererer til idealiserte "speaker - listeners". For å oppnå en fruktbar beskrivelsesmodell trenger man en teori som abstraherer bort aspekter ved språkkompetansen som man ikke ønsker å belyse. Den "omnikompetente" språkbruker i Hymes' kompetansebegrep eksisterer selvsagt ikke som enkeltmenneske, men representerer et teoretisk konstrukt som setter oss i stand til å foreta de ønskede avgrensninger av studieobjektet vårt.

Det som forbinder de to kompetansebegrepene er *kontekstuelle variabler*. Med Lyons kan vi da stille følgende spørsmål: Hvilke typer kunnskap må deltakerne i en kommunikativ situasjon ha, i tillegg til deres kunnskap om språksystemets fonologiske



og grammatiske regler, for å kunne forstå og produsere kontekstuellt vellykkede ytringer?

Dette er bakgrunnen for den tekstanalysen som jeg bruker som dokumentanalyse. Den lar seg beskrive innen rammen av det utvidede kompetansebegrep.

### Tekst og koherens

Jeg skal i det følgende prøve å skissere deler av den tekstlingvistiske semantikken i lys av det kommunikative kompetansebegrepet.

Innenfor iod er det først og fremst Wüsters arbeider, (særlig Wüster 1971) som behandler tematiske enheter. Wüster er opptatt av skillet mellom begrepshierarkier og temahierarkier, som jeg har omtalt, og han redegjør i detalj for de ulike bruksområdene for de to typene, og viktigheten av å holde dem atskilt. Han definerer temaet som "en hver gjenstand som blir behandlet i et dokument" (Wüster 1971:100).

Temabegrepet hans er meget vidt, og refererer til udefinerte størrelser som *Gegenstand* og *Dokument*. Wüster gir ikke noe klart svar på hva disse størrelsene er, og hans temabegrep blir derfor uklart. Han gir heller ingen metodiske retningslinjer for hvordan man kan isolere tematiske størrelser i dokumenter. Hvis vi skal kunne etablere et grunnlag for en slik metode, må vi ta stilling til hva et dokument er. Dette bringer oss naturlig over i spørsmålet om hva en tekst er.

La oss se nærmere på begrepet *tekst* slik det er behandlet innenfor tekstlingvistikken. Både Fossetøl (1979) og van Dijk (1977, 1980) betrakter teksten som en grammatisk størrelse over setningsnivå, som er styrt av den innfødte språkbrukers (kommunikative) kompetanse. Fossetøls verbale ytringer blir da tekstsetninger kun når disse ytringene inngår i en meningsfylt helhet (kalt koherens). Grammatikken utvides og settes i stand til å beskrive teksten som grammatisk fenomen. Det Lyons kaller sosiolingvistik /

stilistikk faller altså innenfor grammatikkens område (jf Lyons 1977).

Jeg skal ikke her begi meg inn på noen videre diskusjon om grensene for grammatikkens beskrivelsesområde, men nøye meg med å konstatere at tekstlingvistikken tilbyr oss en metode for analyse av større enheter enn Chomskys setningsmodell kan tilby. Språkbrukerens kommunikative kompetanse gjør ham altså istand til å skille mellom en tilfeldig sekvens av setninger, og en sekvens av setninger som utgjør en tekst. Koherens blir da en konstituerende del av tekstbegrepet.

Med nøkkelbegreper som koherens og makrostrukturer på den ene siden og tema - rema strukturer på den andre siden, kan man bygge opp en tekstlingvistikk som har til oppgave å beskrive relasjonene mellom de to typene enheter. Johnson-Laird (1987) påpeker at en setning ikke blir lagret som strenger av ord i hjernen, men at vi husker "its sense rather than its syntax."

Johnson-Laird gir her viktige grunner til at vår tekstanalyse bør være semantisk, ikke syntaktisk basert. Johnson-Lairds konklusjon om at vi husker "Its sense rather than its syntax", altså innholdet snarere enn uttrykket, har vel sammenheng med den grunnleggende arbitrære relasjon mellom uttrykk og innhold, som Saussure påpeker. Det er først og fremst dette forholdet som gjør at det tar tid (og krefter) å lære seg fremmedspråk: man må lære seg et nytt sett av (mer eller mindre) arbitrære relasjoner mellom uttrykk og innhold. Men det vi bærer med oss fra morsmålet til fremmedspråket, er først og fremst intuisjonen om at (store deler av) alle språks leksikon lar seg semantisk strukturere i hierarkier, og at bestemte relasjoner (som synonymi, hyponymi, hyperonymi og antonymi) består innenfor og mellom disse hierarkiene. Vi kan vel knapt tenke oss noe språk hvor dette ikke er mulig. Såvidt jeg kan se, er det nettopp dette fenomenet Fossestøl omtaler som det språklige leksikons *infrastruktur*.

Hvis vi ut fra disse betraktningene ønsker å konstituere et indekseringssystem som tar hensyn til disse kunnskapspsykologiske og lingvistiske forhold (ut fra den hypotese at



de kan bidra til å effektivisere systemet), må systemet primært være semantisk, ikke syntaktisk basert. Det er altså tekstens innholdsstruktur som blir vårt primære studieobjekt. Tekstens språklige uttrykksstørrelser blir kun interessante i den grad de fungerer som signaler på hvordan tekstinholdet er organisert (f.eks. tema-remafordeling, bruk av bestemt, ubestemt artikkel, konnektiver og proformer).

Som eksempler på koherensoppbyggende fenomener i tekster nevner Fossestøl anafori, kontiguitet (for eksempel tempusfølge), bestemthet og ubestemthet i nominalfraser og synonymi. Alle disse fenomenene er semantiske relasjoner som strekker seg langt utover setningsgrenser. Sensitivitet overfor koherens betraktes som en del av den omnikompetente språkbrukerens kompetanse. En del av denne koherensen kan forklares ut fra leksikonets infrastruktur (eller tesaurusstruktur).

### **Tema og tekstuell reduksjonsanalyse**

For å nå fra tekst tilbake til term må vi altså gå via tekstuell reduksjonsanalyse til tematiske størrelser, som vil fungere som indeksord og tilordnes deskriptorstatus i en tesaurus. Det finnes flere måter å gjøre dette på. I denne sammenhengen skal jeg gjennomgå to typer reduksjonsanalyse: Tema-remaanalyse og van Dijks reduksjonsanalyse.

Temaet i lingvistikken er først behandlet innenfor Pragskolens tradisjoner, hvor den primært analyseres i rammen av Mathesius' "funksjonelle setningsperspektiv" (FSP). FSP er en måte å organisere ytringer på under en kommunikativ synsvinkel.

Temaet kontrasteres vanligvis med "rema", og dikotomien tema-rema korreleres med skillet mellom "gitt" og "ny" informasjon. Firbas (1964) prøver å beskrive "gitt/ny informasjon" ved hjelp av begrepet "kommunikativ dynamikk". Kommunikativ dynamikk defineres som "den rolle et element spiller i utviklingen av kommunikasjon eller formidlingen av informasjonen". Hvis et element spiller stor rolle i denne

sammenhengen, i den forstand at den "pushes the communication forward" sies den å ha "høy" kommunikativ dynamikk og er ny informasjon, i motsatt fall har den "lav" kommunikativ dynamikk og tilsvarende gitt informasjon (Firbas 1964:270).

Allerede denne tolkningen av temabegrepet reiser spørsmålet om temaet er en diskret størrelse, eller om man kan snakke om grader av temaegenskaper ved språklige elementer i en tekst.

Gjennom det kommunikative dynamikkbegrepet er det imidlertid ikke mulig å etablere en isoleringsprosedyre for tematiske enheter i tekster, som vi jo trenger.

Dyvik (1980) skiller mellom tre typer gitthet: referensiell gitthet, innholdsgitthet og relasjonell gitthet. Den relasjonelle gitthet korrelerer med "gammel" informasjon og inkluderer de to første typene. Den relasjonelle gitthet forklares ut fra det faktum at enheten har "en gitt semantisk rolle i forhold til andre tegn forbundet med seg, eller i det minste gitt som deltaker i en omtalt situasjon" (Dyvik 1980:67). Som eksempel har vi følgende:

#### G

- 1 *Hunden bet postbudet.*
- 2 *Hvem var det som bet postbudet?*
- 3 *Hva gjorde hunden?*

I ytringen "Hunden bet postbudet" som svar på spørsmålet "Hvem var det som bet postbudet?", vil "hunden" være ny informasjon selv om den er referensielt gitt. I "Hunden bet postbudet", som svar på spørsmålet "Hva gjorde hunden?", er "hunden" både referensielt og relasjonelt gitt gjennom sin agensrolletildeling fra verbalet "gjøre", mens både "bitingen" og "postbudet" er ny informasjon. Gjennom å tildele "hunden" agensrollen i spørsmålet, signaliserer avsender at mottaker skal bruke "hunden" som tematisk utgangspunkt i svaret.

Det er altså disse kommunikative signalene mellom avsender og mottaker som



regulerer valget av temaet, mens *hunden* som grammatisk subjekt får sin rolletildeling som agens setningsinternt fra verbet *bite*.

Men denne typen tema er det vi kan kalle tekstsetningstema eller ytringstema. Tekstsetningstemaets hovedfunksjon i tekster er å fastlegge fordelingen av informasjon langs setningssekvenser. Vi trenger også enheter som styrer informasjonsreduksjonsanalysen fra fulltekst til tekstens indekseringsord. Vi er altså ute etter enheter som også organiserer og kategoriserer semantisk informasjon i sekvenser som helhet. Disse enhetene kaller van Dijk *diskurstemaer*. Diskurstemaet definerer van Dijk som en global tekststørrelse som hierarkisk organiserer den tematiske strukturen. Disse diskurstemaene representeres i form av logiske proposisjoner (predikatskalkyler) og kalles *makrostrukturer* hos van Dijk.

Tekstsetningstemaer kontrasteres med remaenheter i tekstsekvenser, noe ikke makrostrukturene gjør. Makrostrukturene angir tekstens hovedobjekt eller hovedidé, og tilsvarer det Brown et al kaller "topic entity" (Brown et al 1986:137).

Det er tekstsetningstemaet som er analyseenheten for Pragskolens tekstlingvister, blant andre Mathesius, Daneš og Červenka. Daneš (1981:42) drøfter teksttemaets betydning for analysen av tekstens oppbygning. Når det gjelder setningselementenes meddelelsesverdi er det remaet som spiller den relevante rollen ettersom den innfører ny informasjon. Når vi derimot ser på den tekstlige strukturingsammenhengen, blir det temaet, med sin lave informasjonsverdi og sin lave kommunikative dynamikk, som kommer til å utføre en viktig strukturingsenhet.

Ut fra dette standpunkt kan enhver tekst (og ethvert tekstavsnitt) betraktes som en temasekvens som kan konstateres gjennom såkalte kompletteringsspørsmål.

Kompletteringsspørsmål påstås av Daneš å være et entydig kriterium for å bestemme tema og rema i en tekstsetning. Et utsagns kompletteringsspørsmål gir remaet, mens resten blir tema. I setningen "Han fikk en bok av en kollega", med kompletterings-

spørsmålet "Hvem fikk han boken av?", gir remaet "av en kollega". "Han fikk en bok" blir da temaet.

Daneš fremholder videre at en og samme (system)setning kan utfylles med mange forskjellige kompletteringsspørsmål (som postbudeksempelen i G illustrerer). Dette forklares som setningens skiftende "meddelelsesperspektiver", variabler som styres av kontekster og situasjoner. Det vil i praksis si at en og samme systemsetning representerer et (muligens) infinitt antall ytringer.

Den tematiske strukturen i en tekst sees på som en tematisk sammenlenking med forbindelsesmuligheter mellom temaene. Den tematiske progresjon sees både i relasjon til tekstavsnitt og tekster i sin helhet. Hele komplekset kaller Daneš *temafølge* (TF).

Daneš stiller spørsmål om hvilke seleksjonskriterier man skal legge til grunn for temaer i en utsagnsfølge. Hvilke relasjoner forekommer mellom tekststykkets temaer, og i hvilken relasjon står disse temaene til det såkalte "hypertemaet" i tekststykket? (Pragskolens "hypertema" synes i praksis å overlape med van Dijks "makrostruktur", men begrepene synes å være teoretisk forskjellig forankret). Dette er i grove trekk utgangspunktet for Daneš' analyse av teknisk og annen faglitteratur på tjekkisk.

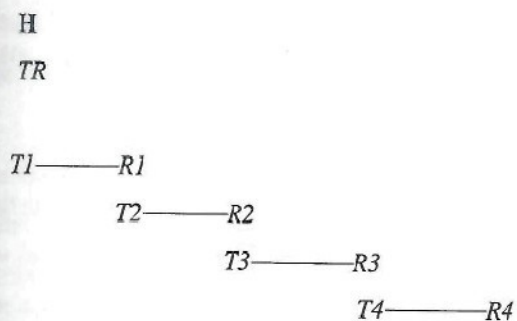
En setnings informasjoninnhold består av selve utsagnet pluss den mengden av konklusjoner en kan trekke av den og de foregående setningene. Et tema i en ny setning virker da som en seleksjonsmekanisme. Tematisering vil da si å velge ut av en setning eller av den setningens konklusjonspotensial et bestemt element (eller flere) som da blir setningens kontaktsrelasjon til det foregående. Fra dette "gitte" elementet kan teksten ekspandere i ulike retninger avhengig av det budskap senderen vil formidle. (Fossestøl 1979:367f).

Den tematiske progresjonen fremstilles da som en sekvens av (rema - (tema + rema) - (tema + rema) - (tema + rema)) etc.. Det vil si at remaet i en tekstsetning blir repetert i form av tema i en ny tekstsetning. Resultatet blir en *temafølgestructur* (TF - struktur).



Červenka (1981:50ff) angir to hovedtyper av slike TF-strukturer (jf H og J). I H er et tema (T) i et etterfølgende utsagn forbundet med et rema (R) i et forutgående utsagn. Denne typen kalles TR. I J er et tema i etterfølgende utsagn forbundet med et tema (T) i et foregående utsagn. Denne typen kalles TT.

Ved TR realiseres tekstens koherens langs den syntagmatiske akse (dvs sekvensielt i teksten):



Men tekster kan inneholde samme tematiske struktur som i H og likevel føles inkoherente:

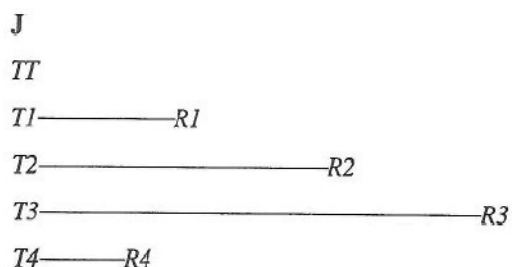
I

*Jeg bor i Fana i Bergen. Bergen er Norges nest største by. Norge ligger i Skandinavia. Skandinavia ligger i Nordeuropa. Nordeuropa er en del av Europa. Europa er en verdensdel, etc.*

Tekst I mangler et "poeng" og vi blir fristet til å spørre etter det, hvis ingen påfølgende setninger binder tematikken tilbake til første setning. Vi trenger altså en type global struktur som strekker seg lengre enn over to tekstsetninger.

I tekster av typen TT, derimot, fyller utsagnenes remaer ingen koherensdannende

funksjon. Forbindelser etableres gjennom gjentakelser og repeterte tilknytninger til et felles utgangspunkt. Tekstens kommunikative dynamikk beveger seg fra ett tema (eller ett aspekt av et felles hypertema) til et annet tema. Tekstkoherensen sies da å være realisert hovedsakelig langs den paradigmatiske aksen, jf J:



Et "hypertema" er i følge Daneš en temafølge med gjennomgående tema, eller et implisitt tema i en tekst som andre temaer kan avledes fra.

I indekseringssammenheng synes det klart at T-enhetene, særlig i J er klare kandidater til emneindekseringsord. Et "hypertema" (i den grad det lar seg utkrystallisere) vil da eventuelt få større vekt (f. eks. T1-T4 i J, hvor temaet er gjennomløpende, i motsetning til T1-T4 i H). *Hypertema* og *makrostruktur* kan, som jeg sa, overlappe, men kan også være forskjellige. Isoleringsprosedyrene synes i alle fall å være ulike. Van Dijk går ut fra helheten og reduserer ved hjelp av reduksjonsregler, mens Daneš og Červenka går direkte på sekvenser av setninger i tekster uten først å ta hensyn til helheten, den globale strukturen. Men utover kompletteringsprinsippet angir ikke Daneš og Červenka noen isoleringsprosedyrer for tematiske enheter.

Van Dijk identifiserer diskurstemaer med det han kaller "makrostrukturer" som danner nivåer eller hierarkier.

For å kunne isolere makrostrukturer må vi ifølge van Dijk foreta en informasjonsreduksjonsanalyse. Til dette formål etablerer van Dijk en rekke reduksjonsregler som alle bygger på implikasjon ("entailment").



Men hvis man som van Dijk tar utgangspunkt i helheten, bør man ha en idé om hva helheten er. Man kan stille seg spørsmålet om hva en fullstendig beskrivelse av noe er.

Som jeg sa med referanse til Greimas realiseres leksemet aldri som helhet i en tekst, men gis bare en partiell realisering. Dette har sammenheng med det faktum at enhver tekst representerer et utvalg av den totale mengden mulig informasjon. Forfatteren foretar derfor et utvalg av totalmengden. Vanlig tekst inneholder med andre ord bare de fakta som er *pragmatisk relevante*.

Vi har altså å gjøre med partielle beskrivelser av objekter, fenomener og tilstander. Vår fulltekst eller utgangstekst (den teksten vi skal indeksere) er altså et tematisk utvalg av en totalitet, og man antar at den partielle beskrivelsen er kompatibel med helheten, den "totale" beskrivelsen.

Det er klart at avsender ønsker å informere, men han kan også formidle andre aspekter enn det informative. Språket er et komplekst samspill mellom distinktivitet og redundans, dvs. informasjonsformidling og redundante understøttelsesmekanismer som har som funksjon å lette overføringen av informasjonen i kommunikasjonen. Jo mindre redundansen i kommunikasjonssituasjonen er, desto større blir den kommunikative belastning for mottakeren, som jeg har vært inne på.

Faktorer som koherens, temavalg og makrostrukturer har som en viktig funksjon å redusere den kommunikative belastningen. Ifølge van Dijk har makrostrukturene til oppgave å begrense mengden av mulige begreper, og fungerer som en global begrensning på leksikalske valg. En doktoravhandling om elementærpartikler i fysikken, for eksempel, vil neppe inneholde begreper som "kjærlighet", "Gud", og "mord".

Ut fra vår kommunikative kompetanse (vår kunnskap om språket og verden) er vi i stand til å strukturere en tekst i et nettverk med temaenheter som har innbyrdes relasjoner. Vi tolker inn temaene i semantiske felt som overensstemmer med vårt

kunnskapsbilde.

## Makrostrukturer og reduksjonsregler

Makrostrukturer kan formuleres språklig som proposisjoner eller påstander i setningsform i naturlig språk. Hvis en proposisjon er sammensatt av flere proposisjoner, gjelder også kriteriet om at de må være implisert av tekstsekvensen som helhet. På denne måten får man frem deltemaer ("subtopics" eller "atomic topics") som danner implikasjonelle hierarkier.

Det overordnede temaet i en tekst er da den proposisjonen som ikke er implisert av noen andre i teksten. Det overordnede temaet er tekstens mest globale makrostruktur. Det er svært mange interessante spørsmål man kan stille i forbindelse med denne teorien, men vi har ikke tid til å gå inn på dem er.

I det følgende skal vi kort beskrive van Dijks reduksjonsregler, dvs hans makroregler.

## Makroreglene

Hovedpoenget med van Dijks makroregler er å genere resyméer fra utgangstekster. For å kunne gjøre dette må utgangsteksten først omformes til en underliggende proposisjonal form, dvs en påstandsform hvor det denotative innholdet er bibeholdt, mens ekspressivt og sosialt innhold er "silt vekk". Påstandene i denne underliggende formen kalles mikroproposisjoner.

De proposisjonene som fremkommer etter at reduksjonsreglene har virket, kalles makroproposisjoner. Makroproposisjonene danner resyméets underliggende form:



## K

*Utgangstekst*

*resymé*

*mikroproposisjoner-----> reduksjonsregler-----> makroproposisjoner*

Makroreglenes oppgave er å definere en teksts globale tema. Siden reglene opererer på mikroproposisjoner og gir makroproposisjoner, vil reglenes utdata ta form av resyméer.

Generelt kan reglene beskrives som reduktive, fordi de reduserer informasjonsinnhold i tekster. En av reglene har konstruktivistiske egenskaper, fordi den konstruerer ny informasjon. Alle reglene er organisatoriske, fordi de kombinerer mikroproposisjoner til makroproposisjoner. Det er fire makroregler (jf van Dijk 1980:82f): DELETION, GENERALIZATION, CONSTRUCTION og ZERO.

DELETION stryker alle de proposisjonene i tekstbasen som ikke er relevante for tolkningen av andre proposisjoner i diskursen, og som ikke denoterer saksforhold som kan subsumeres som normale egenskaper av et mer globalt saksforhold som er denotert av tekstens makroproposisjon. DELETION stryker altså all informasjon som ikke er nødvendig for å forstå hovedidéen i en teksts budskap. Som eksempel kan vi ta følgende:

## L

- 1 *Jon kunne ikke komme på møtet.*
- 2 *Han var syk.*
- 3 *Møtet måtte utsettes.*

Gitt en passende kontekst vil følgende implikasjonsforhold være gjeldende: I L er 2 implisert av 1, og kan strykes ved DELETION. Videre er 1 implisert av 3 og kan strykes.

Ved GENERALIZATION er det ikke snakk om stryking av globalt irrelevante

proposisjoner, men abstraksjon fra semantiske detaljer i de respektive setningene ved å konstruere en proposisjon som konseptuelt sett er mer generell. De respektive individuelle deltakerne (på lokalt nivå) kan grupperes, mens predikatet i de respektive proposisjonene subsumeres under en fellesnevner som denoterer "the superset" av den egenskapen eller relasjonen som denoteres. Et eksempel er M:

M

- 1 *Per spiller fotball, Kari hopper tau og Lars blåser i fløyte.*
- 2 *GENERALIZATION: De leker.*

CONSTRUCTION er mer kontroversiell. Den bygger blant annet på Roger Schanks rammebegrep og Hilary Putnams stereotypibegrep.

I denne regelen syntetiseres mikroproposisjonene ved å erstatte dem, som en sammenbundet sekvens, av en proposisjon som denoterer normale komponenter, betingelser eller konsekvenser, som i eksempel N:

N

- 1 *Jon var syk. Han ringte etter lege.*
- 2 *Jeg reiste til New York.*

I sekvens 1. i N (Jon var syk. Han ringte etter lege.) definerer sekvensen av proposisjoner makroproposisjonen som helhet. Makroproposisjonen denoterer en mer eller mindre stereotypisk sekvens av hendelser i en episode hvor det er konvensjonelt kjent hvilke egenskaper og saksforhold som vanligvis assosieres med den.

Ved CONSTRUCTION kan lokale proposisjoner være relevant for tolkningen av en annen proposisjon i teksten. Man kan for eksempel konstruere proposisjonen (2. i M) "Jeg reiste med fly til New York" som en makroproposisjon som globalt representerer en kompleks handling hvor alle typer detaljer er involvert. For eksempel det å reise til



flyplassen og sjekke inn, få bagasjen sikkerhetsjekket, etc.

Noen av detaljene (for eksempel det å få et ombordstigningskort) *kan* være relevant for det faktum at ombordstigningskortet leveres til en luftfartsoffiser senere i sekvensen. Det samme kan for eksempel gjelde for bagasjeetikett.

Slike proposisjoner kan være *sekvensielt relevante* men ikke på et mer globalt nivå med mindre proposisjonen bestemmer tolkningen av senere proposisjoner som selv kan være identiske med en makroproposisjon. Den globale proposisjonen kan bare konstrueres på grunnlag av konvensjonelt kjente aspekter ved den globale hendelsen. Det vil si på basis av proposisjoner som opptrer i en kunnskapsramme.

Tekstuelle proposisjoner som ikke er konvensjonelt kjent (for eksempel det å møte en bekjent på flyplassen) vil ikke kunne håndteres av denne regelen. I så fall er slike proposisjoner enten makroproposisjoner (ved at de ikke reduseres av CONSTRUCTION eller GENERALIZATION) eller de er detaljer ved en flyreise, og strykes av DELETION eller GENERALIZATION.

Det er imidlertid viktig å presisere at det som denoterer normale betingelser, komponenter eller konsekvenser av fenomener, kan være annerledes i et spesielt fagområde enn i allmennspråket. Sett av stereotyper innen et fagområde må derfor identifiseres og presiseres hvis man skal kunne anvende regelen på tekster fra dette området.

Van Dijk har ikke gitt noen eksplisitt formulering av ZERO-regelen, men påpeker at vi trenger en regel som sikrer at visse mikroproposisjoner forblir intakte på makronivå. ZERO-regelen tar en mikroproposisjon som inndata, og gir samme proposisjon som utdata på makronivå (som for eksempel ved tematiske proposisjoner).

Disse reduksjonsreglene tar oss altså fra fulltekst til resymé. Hvordan den videre veien fra resyméer til deskriptorer skal finnes er fremdeles et vanskelig spørsmål. Vi tar da

et sprang fra sannhetsbetinget innhold til uttrykk som refererer til enheter i verden, referenter.

### 3 Konklusjon

Denne rundturen har tatt oss fra termnivå til tekstnivå ved hjelp av definisjoner og forklaringer. Det aktualiserte økonomiaspektet i informasjonsteorien. Populariseringsaspektet ble tatt frem som et godt eksempel på formidling mellom nivåene. I den andre delen har vi prøvd å gå motsatt vei, fra tekst til term. Denne veien er det vanskeligere å gå, blant annet fordi utgangspunktet, helheten er mye mer komplisert og vanskeligere å definere/avgrense. Verken tema-remadistribusjon på yringsinternt nivå eller makrostrukturell reduksjonsanalyse er tilstrekkelige til å vise forholdene mellom de to nivåene.

### Litteratur

- Andersen, Ø. (1989). Fagspråk som forskningsdisiplin (1. del)  
Bergenholtz, H & Molly, H. (red.) *Hermes 3. Tidsskrift for Sprogforskning*. Det Erhvervsproglige Fakultet. Handelshøjskolen i Århus.
- Andersen, Ø. (1990). Fagspråk som forskningsdisiplin (2. del)  
Bergenholtz, H. & Molly, H. (red.), *Hermes 4. Tidsskrift for Sprogforskning*. Det Erhvervsproglige Fakultet. Handelshøjskolen i Århus.
- de Beaugrande, R. & Dressler, W. (1992). *Introduction to Text Linguistics*. Singapore: Longman.
- Brown, G. & Yule, G. (1986). *Discourse Analysis*. Cambridge: Cambridge University Press.
- Daneš, F. (1981). Om en lingvistisk analys av textens struktur.  
de Kaminsky, J. & Laven, G. (red.) *Textkoherens. Nya riktninger i studiet av textens struktur*. Universitetet i Uppsala: Uppsala.
- Dyvik, H.J.J. (1981). Om definisjonen av "tema" i tema/remadistribusjonen. *Egenproduksjon*. 8. Nordisk institutt. Universitetet i Bergen.
- Firbas, J. (1964). On defining the Theme in Functional Sentence Analysis. *Travaux de linguistique de Prague*. 1. Prague.
- Fossestøl, B. (1980). *Tekst og Tekststruktur. Veier og Mål i tekstlingvistikken*. Oslo: Universitetsforlaget.
- Grice, P. (1975). Logic and Conversation. Cole, P & Morgan, J.L. (red.) *Syntax and Semantics*. New York: Academic Press.
- Greimas, A.J. (1981). Aktanter, aktører och figurer. de Kaminsky, J. & Laven, G.



- (red.) *Textkoherens. Nya riktningar i studiet av textens struktur*. Universitetet i Uppsala: Uppsala.
- Hymes, D. (1971). *On Communicative competence*. Philadelphia: University of Pennsylvania Press.
- Johnson-Laird, P. N. (1987). *Mental Models*. Cambridge: Cambridge University Press.
- Knudsen, T og A. Sommerfelt (1957). Principles of Unilingual Dictionary Definitions. *Reports for the Eight International Congress of Linguistics*. Oslo: Oslo University Press.
- Kuhn, Th. (1970). *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.
- Lang, F.H. (1980). Inhaltserschliessung. Laisiepen, K, Lutterbeck E.& Meyer-Uhlenried K-H. (red.). *Grundlagen der praktischen Information und Dokumentation*. München. DGD-Schriftenreihe. Band 1: K.G.Saur Verlag.
- Laurén, Chr. (1985). Definition och forklaring. *Nordisk Terminologikursus*. II. Skodsborg. Manuskript.
- Lyons, J. (1977). *Semantics*. I og II. Cambridge: Cambridge University Press.
- Martinet, A. (1955). *Economie de Changements phonétiques*. Bern: Francke.
- Picht, H. & Draskau, J. (1985). *Terminology. An Introduction*. Guildford: University of Surrey Press.
- Platzack, Chr. (1977). Substantivsjuka - ett friskhetstecken? *Lundastudier i nordisk språkvetenskap*. D 12. Lund. Lund 1973-77.
- Roald, J. (1985). termen - et faglig begreps uttrykksside. *Nordisk Terminologikursus*. II. Skodsborg. Manuskript.
- Roald, J. (1986). Om begrepet fagspråk. *Romansk fra vest*. Skriftsserie. Romansk institutt, Universitetet i Bergen.
- Sparck Jones, K & Kay, M. (red.)(1973): *Linguistics and Information Science*. New York: Academic Press.
- Tauli, V. (1968). Introduction to a Theory of Language Planning. i: *Acta Universitatis Upsaliensis*. Uppsala. Universitetet i Uppsala.
- Wüster, E. (1971). Begriffs- und Themaklassifikationen. Unterschiede in ihrem Wesen und in ihrer Anwendung. *Nachrichten für Dokumentation*. 2. Frankfurt/M.