

Øivin Andersen

Argumentstruktur og nominalisering

In this article the main focus is on theoretical and representational issues connected to argument structure in nominalizations. The relevance of argument structure theory for lexicographical and terminological analysis will be advocated and exemplified by nominalizations of technical terms in documentation related to the Norwegian offshore petroleum industry. Representational problems connected to the systematic ambiguity of Norwegian verbal nouns denoting events (Grimshaw 1990) will be discussed and illustrated by examples from the Norwegian data base NOT.

1. Bakgrunn

I generell lingvistikk har det i lang tid vært tradisjon å skille mellom teoretisk (eller allmenn) lingvistikk på den ene siden og anvendt lingvistikk på den andre siden (på en måte analogt med skillet mellom leksikologi og leksikografi innen studiet av menneskelige språks leksikon, og skillet mellom terminologi og terminografi innen den allmenne terminologilæren i Wüsters tradisjon).

Disse skillene er problematiske og ofte gjenstand for både kritikk og nye fortolkninger. Leksikografi har av mange tradisjonelt vært ansett for å være en praktisk disiplin som teoretisk anlagte lingvister befattet seg med i minst mulig grad. Den ble ansett for å være en eller annen form for anvendt lingvistikk, til tross for at teoretiske studier av leksikon-strukturering i mange miljøer innen generell lingvistikk overhodet ikke stod på dagsorden. Når allmennlingvistene viste så liten interesse for den allmenne leksikografien er det lett å forstå at den fagspråklige leksikografien, og ikke minst den Wüster-inspirerte terminologien, nærmest var ikke-eksisterende innen disse miljøene.

Det som først og fremst har endret lingvistenes syn på leksikons plass i syntaktisk beskrivelse, er det jeg vil kalle komputasjonaliseringen av allmenn lingvistikk. Dette har skaffet fotfeste for fremveksten av en ny disiplin: datalingvistikk. I første omgang ble datalingvistikk avfeid og henvist til en plass blant de andre anvendte "bindestreksdisiplinene", men etter hvert som den teknologiske fremgang har stormet videre, har datalingvistikken rykket inn i hjertet av det som tradisjonelt har vært allmennlingvistenes domene: syntaks. Det området av datalingvistikken som kanskje har lengst tradisjon, er maskinoversettelse, som man eksperimenterte med allerede på 60-tallet.

I flere år syslet allmennsyntaktikerne med regelsystemer på data-maskinene sine med oppkonstruerte og kunstige eksempler. Dette viser også igjen i det som ble skrevet innenfor området. Setninger som: "The cat is on the mat", "John kicked the ball" og "The farmer killed the duckling", gikk igjen og igjen. Dette viser tydelig at leksikon var forsømt.

Man kan saktens spørre hvorfor. Tidligere var fokus på syntaks. Syntaksen var autonom, dvs. prinsipielt uavhengig av faktorer i leksikon. Senere, dvs. på 80-tallet og utover på 90-tallet, kom andre syntaktiske teorier enn Chomskys generative syntaks i ulike varianter på banen: Bresnans LFG-syntaks (leksikalsk funksjonell grammatikk), for eksempel, legger større vekt på leksikons betydning for syntaktisk analyse. I denne forbindelse kan vi også nevne skillet mellom syntaktiske teorier som setter setningen i fokus, til forskjell fra andre som setter verbalet i fokus som kjernen i setningen (dvs. avhengighetsgrammatikker, jf Tesnière, Helbig). Det var først og fremst den sistnevnte typen syntaksteorier som fokuserte på argumentstrukturer forbundet med setningens hovedverb (jf begrepet valens).

I stigende grad har man måttet innse at for å få automatisk syntaktisk analyse til å fungere, må de syntaktiske reglene, prinsippene og parametrene forbindes med et komputasjonelt, strukturert leksikon av større omfang (cf Pustejovsky 1991).

Man oppdaget også etter hvert at komputasjonelle grammatikker koplet med et komputasjonelt kompatibelt leksikon er vesentlig for utforskningen av lingvistiske strukturer i menneskelige naturlige språk.

I tillegg har den tidligere så utskjelte korpuslingvistikken fått en renessans, blant annet fordi datamaskinenes lagringskapasitet har økt betraktelig i de senere år. Dette, sammen med utvikling av avanserte tekstkodningssystemer som HTML, har ført til at den modulære kombinasjonen av elektroniske grammatikker og elektroniske leksika kombinert med korpuser er særdeles viktige redskaper for lingvistisk forskning.

I tillegg til dette kan vi på den fagspråklige siden si at de komputasjonelle modulpakkene kanskje har større sjanse for å bli funksjonelle enn de allmennspråklige, fordi den leksikalske struktureringen innen et gitt fagspråklig domene vil være lettere avgrensbar og dekomponerbar enn allmennleksikalske strukturer. Veien fra eksperimentell produktutvikling til kommersielle produkter synes å være kortere innen en fagspråklig kontekst enn innen en allmennspråklig. Dermed er det selvsagt ikke sagt at veien er kort og rett frem.

Komputasjonaliseringen og fremveksten av datalingvistikk som egen disiplin har altså økt behovene for leksika tilrettelagt for lingvistisk forskning. Dette gjelder både flerspråklige leksika (til bruk for

eksempel i eksperimentelle maskinoversettelsesprosjekter) og enspråklige (for eksempel til bruk for lingvistiske analyser av enkeltspråk). Det fagspråklige domenet er spesielt vel tilrettelagt for dette av flere grunner: Det er større grad av isomorfi mellom uttrykk og innhold i terminologiske data enn i allmennleksikografiske data, og de pragmatiske fortolkningsrammene er mer velavgrenset enn i allmennspråket, selv om det er snakk om gradforskjeller her).

I denne artikkelen skal jeg ikke berøre alle de komputasjonelle og lingvistisk-leksikografiske problemene som er forbundet med komputasjonelle leksika, men fokusere på et felt som jeg mener har interesse både for fagleksikografien og terminologien, uansett hvilken tradisjon man hører til innenfor disse områdene: argumentstrukturer assosiert med verbalnominaliseringer i en bestemt type teknisk dokumentasjon: driftshåndbøkene for Gullfaks A-plattformen i Nordsjøen.

2. Argumentstruktur

Begrepet **argumentstruktur** kalles blant annet også **subkategoriseringsstruktur** (Chomsky) og **valensstruktur** (Tesnière). Dessuten har de nevnte metatermene **semantiske roller** og **argumentstrukturer** blitt brukt om hverandre som nærsynonymer.

Denne metaterminologiske mangfoldighet er ikke bare resultat av forskjeller i teoretisk forankring, men illustrerer også den kompleksitet fenomenet oppviser i menneskelige naturlige språk.

I enhver analyse og beskrivelse av disse fenomenene er det viktig å skille mellom syntaks og semantikk. Termer som **argumentstruktur** og **subkategorisering** vil derfor betraktes som syntaktisk-leksikalske uttryksrelaterte termer, mens deltakerrolle er en semantisk leksikalsk term. **Argumentstruktur** og **deltakerrolle** er dermed å betrakte som leksikalske fenomener som har vidtgående syntaktiske virkninger.

Teorier om **semantiske roller** har også opptatt både lingvister og filologer i lang tid, men tilnæringsmåtene og bruken av teoretiske begreper om emnet har også på dette området stor variasjonsbredde. Semantiske roller har også vært kalt **semantisk kasus**, **dypstrukturell kasus** (Fillmore), **deltakerrolle** (Grimshaw), **aktant** (Tesnière) og **thetarolle** (Chomsky).

Generaliseringer om argumentstruktur har tidligere vært ansett for å være en syntaktisk størrelse hvor fenomener som transitivitet og passiv var relevante beskrivelsesområder.

I senere lingvistiske arbeider behandles argumentstruktur som et leksikalsk fenomen (Grimshaw 1990). Det har med leksikalsk informasjon om et leksems grammatiske egenskaper å gjøre, og hører like

naturlig hjemme i en leksikonrepresentasjon som morfologiske opplysninger som numerus og genus ved substantiver.

Dette aktualiserer et klassisk problem i lingvistik: Grenseopp-gangene mellom grammatikk og leksikon. I et generelt, allmennspråklig langueperspektiv synes grensene å være uklare med glidende over-ganger. Dette virker også intuitivt plausibelt når man kopler grenseopp-gangen med begrepene **forutsigbarhet** og **generaliserbarhet**. Forutsigbarhet er klart et graduert begrep, og fortolkes slik at høy grad av forutsigbarhet knyttet til et lingvistisk fenomen er en indikator på dets grammatiske status, og at fenomenet prinsipielt lar seg redusere til produktive (generaliserbare) regler.

Et klassisk eksempel er regelmessig (svak) bøyning av verber i for eksempel norsk. Ved disse verbene kan man foreta generaliseringer over klasser, mens de uregelmessige (sterke) verbene ikke i samme grad lar seg generalisere over. Likevel har de sterke verbene et preg av en slags regelmessighet, fordi de ofte utgjør de vanligste verbene i et kjerneleksikon. Den logiske konsekvensen er da å markere de uregelmessige i leksikon, mens de regelmessige får sine generaliseringer nedfelt i en grammatikk.

Ved nærmere studium av disse to typene verber viser det seg snart at generaliseringene i klasser ved svake verb ikke er så forutsigbare som man ønsker å tro, og at det ved de sterke verbene er mulig å foreta generaliseringer som kan formuleres som delproduktive regler.

Om grensene mellom grammatikk og leksikon "egentlig" er veldefinerte eller ikke, er vanskelig å konstatere. Det er nok i høy grad et spørsmål om teoretisk ståsted. De som har en veldefinert grenseopp-gang som utgangspunkt, vil hevde at grunnen til at overgangene ser glidende ut, er at vi ennå ikke har studert fenomenene tilstrekkelig nøye, og ennå ikke funnet de mest fruktbare og hensiktsmessige redskapene til formålet.

La oss forsøke å anskueliggjøre argumentstrukturbeskrivelsens plass i dette bildet ved å etablere en enkel arbeidsmodell hvor grammatikken plasseres på det ene ytterpunktet i en linje og leksikon på det andre ytterpunktet:

Modell I

grammatikk.....leksikon

I denne modellen vil det være naturlig å plassere argumentstrukturbeskrivelsene nær den leksikalske polen fordi argumentenes antall og realisering ikke uten videre lar seg generalisere over på tvers av det

enkelte leksem. Analyse av de assosierte deltakerrollene tyder også på at det enkelte leksem som sådant er det naturlige beskrivelsesdomenet.

Men det er også en del konsekvenser av argumentstrukturfenomenet som har vidtgående syntaktiske konsekvenser, som vi har nevnt. I tillegg viser den leksikalske forskningen at tradisjonelt leksikografiske fenomener som polysemi har visse generaliserbare egenskaper som peker i retning av den grammatiske pol (jf Pustejovsky et al. 1996a). Noen av disse polysemifenomene lar seg disambiguere ved argumentstrukturanalyse. Dette tar jeg opp mer detaljert i forbindelse med Grimshaws analyse av verbalnominaler, som jeg anvender i analysen av nominaliserte termer i NOT-basen.

Det er også nødvendig å skille mellom argumentstrukturangivelse og **kollokasjonsangivelse**. Ifølge Benson (1989) betegnes kollokasjoner som "arbitrary recurrent word combinations". Dette tyder på at kollokasjonsangivelser ligger nærmere den leksikalske polen enn argumentstrukturangivelser. Pustejovsky (1991:422) ser på kollokasjoner som toveisrelasjoner (kalt **kospesifikasjoner**). Ved argumentstrukturtildeling er det snakk om en enveis relasjon fra verb/ verbalnominal til argumentene, mens ved kospesifikasjon vil verb/ verbalnominal riktignok tildele argumentstrukturrammer, men argumentet selv er også i stand til å selekttere type verb som skal styre det. Derfor vil, i Pustejovskys analyse, forskjellen mellom *bake a potato* og *bake a cake* ikke være polysemi ved leksetet *bake*, men en forskjell i kospesifikasjonsforhold mellom henholdsvis *bake* og *a potato* (naturprodukt) og *bake* og *a cake* (artefakt).

Det synes også å være en forskjell i generaliseringsnivå: Argumentstrukturer angir typiske realiseringer (dvs. hovedtyper), mens kollokasjoner angir mer detaljert stilistiske varianter som i større grad er bundet til spesifikke kontekster. Det synes naturlig å se på kollokasjoner som en presisering av argumentstrukturene i en videre leksikalsk utbygging.

Den grunnleggende, sentrale ideen i argumentstrukturteori er at verbet i predikatsfunksjon i setninger har en slags pivotalstatus hvor den tilskrives en rolletildelingsfunksjon. På det syntaktiske plan består denne funksjonen i tildeling av obligatoriske (fortrinnsvis) nominale ledd som er konstituenten av setningen eller setningens predikat, og som dermed utfører syntaktiske funksjoner som subjekt, direkte objekt og indirekte objekt. På det semantiske plan skjer det en parallell tildeling av semantiske deltakerroller som argumentleddene er bærere av: agent, patient, experiencer, theme, source, goal, etc.

Ad hoc-pregede motsetninger mellom ulike tilnæringsmåter oppstår særlig på det semantiske plan: Valget av semantiske etiketter og deltakerrollenes definisjoner og antall varierer mye, og lenkingen

mellom syntaktiske argumenter og semantiske roller har vært analysert på varierende måter. Dowty (1989:70) oppsummerer tilstanden slik:

In fact, two problems with thematic roles remain persistently vexing: (i) lack of agreement among linguists as to which thematic roles exist, and the absence of any obvious way to decide this question, and (ii) the lack of any effective way to independently justify the assignment of noun phrases to thematic roles in particular sentences.

Når det gjelder forholdet mellom syntaktisk og semantisk tildeling, var utgangspunktet i den fillmorske kasusgrammatikken semantiske dypstrukturelle kasus (dvs. deltakerroller). Det var dem som bestemte de syntaktiske overflateargumentene. Dette var et generelt trekk ved den generative semantiske tradisjonen.

Senere varianter som Rappaport (1983), Grimshaw (1990) og Dowty (1989) legger vekt på at semantiske roller "indekserer" predikaters argumenter semantisk på en parallell måte som, men prinsipielt uavhengig av, den syntaktiske kodingen av grammatiske setningsfunksjoner som subjekt og direkte objekt. Hos Rappaport (op.cit.) koples disse to nivåene sammen ved hjelp av lenkingskonvensjoner.

Når det gjelder argumentstruktur ved verb, synes det i den moderne litteraturen å være allmenn enighet om at tildeling skjer både på det syntaktiske og det semantiske plan. Uenigheten er imidlertid større når det gjelder tildeling ved verbalnominaler. Dowty (op.cit.:90) konkluderer med at nominaler som beskriver en hendelse, ikke har subkategoriseringer, slik verber har:

1. *Parkeringen av bilen foran inngangen førte til kaos.*
2. *Parkeringen av bilen førte til kaos.*
3. *Parkeringen førte til kaos.*

Et slikt resonnement baserer seg tydeligvis på et skille mellom det dependensgrammatikerne kaller valensbundne og valensuavhengige ledd. Det betyr i så fall at det korresponderende verbet *parkere* "krever" tre valensbundne argumenter: *han, bilen og foran inngangen* i:

4. *Han parkerte bilen foran inngangen.*

Valensbinding ved argumenter blir vanskeligere å fortolke når setninger som 5, 6 og 7 umiddelbart synes å være like akseptable som 1, 2 og 3:

5. *Han parkerte bilen.*
6. *Han parkerte foran inngangen.*
7. *Han parkerte.*

Det synes derfor åpenbart at rolletildelingsrammene både for verber og for verbalnominaler er underlagt begrensninger som er langt mer komplekse enn valensbinding tilsier. Blant annet synes pragmatiske faktorer som implisitet i kontekst å komplisere dette bildet.

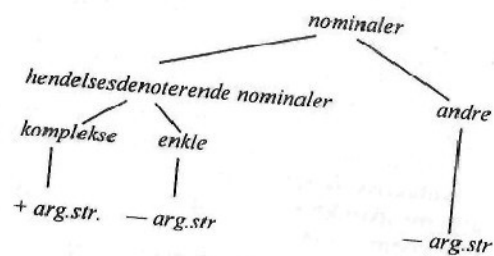
Rappaport (op.cit.:113) hevder at forholdet mellom et verb og dets korresponderende verbalnominal skal representeres med et felles sett av argumentstrukturer slik at for eksempel det leksikalske oppslaget for *parkere* og *parkering* vil ha de samme opplysninger om argumentstruktur. Rappaport ser på argumentstrukturer som en dypstrukturell enhet som realiseres syntaktisk og/eller semantisk på overflatenivå. Poenget er at felles argumentstruktur ikke impliserer felles syntaktisk struktur. Et annet viktig poeng (som går tilbake til Chomsky 1970) er at de syntaktiske realiseringene av argumenter er rikere ved verber enn ved verbalnominaler i den forstand at visse realiseringer som *subjektsløfting*, *"dative shift"*, etc. er blokkert ved nominalisering.

I motsetning til Dowty hevder Grimshaw (1990) at visse nominaler som denoterer en hendelse, har argumentstruktur. Skillet mellom de nominalene som har argumentstruktur og de som ikke har det, kan forklares denotasjonelt og bidrar dermed til å forklare en systematisk ambiguitet ved denne type nominaler. Tildelingen av både argumentstruktur og assosierte deltagerroller skjer parallelt i leksikon ved verber og deres tilsvarende verbalnominaler. En viktig forskjell er at nominaler ikke kan tilskrive sine argumenter kasus direkte, men indirekte via preposisjoner. Dette forklarer forskjeller som for eksempel *"dative shift"* i 9, men ikke i 11:

8. *Kongen overrakte en medalje til ham.*
9. *Kongen overrakte ham en medalje.*
10. *Kongens overrekkelse av en medalje til ham.*
11. **Kongens overrekkelse ham av en medalje.*

Som nevnt lar verbalnominaler seg ifølge Grimshaw disambiguere ved argumentstruktur: De nominalene som tar argumentstruktur, lar seg definere denotasjonelt: De nominaler som denoterer hendelser, kan deles i dem som denoterer komplekse hendelser og dem som denoterer enkle hendelser. De førstnevnte har en obligatorisk argumentstruktur, mens de sistnevnte ikke har det. La oss sette det opp i en oversikt:

Modell II



Denne distinksjonen korresponderer med den mer velkjente prosessen/produktlesingen ved verbalnominaler, for eksempel *arkitektens tegning av plattformen*. Ved en prosesslesing (komplekshendelsesdenoterende) refereres det til den komplekse hendelsen "det å tegne en plattform", mens det ved resultatlesingen refereres til resultatet av denne prosessen.

Ved prosesslesingen blir både *arkitektens* og *av plattformen* valensbundne ledd som får sin semantiske rolle tildelt av nominalet, mens de to leddene i en produktlesing ikke får tildelt rolle, selv om de er underlagt seleksjonsrestriksjoner. Ved prosesslesingen får det premodifiserende leddet en agentiv rolle, ved produktlesingen uttrykker det premodifiserende leddet en slags tilhørighetsrelasjon som er ganske vag og underlagt pragmatiske betingelser.

De to lesingene får også konsekvenser for de ulike leddenes status. Det premodifiserende leddet blir nominale ledd med argumentstatus og rolletildeling fra nominalet, mens de blir modifikasjonsledd uten slik status ved andre lesninger.

Lødrup (1989:139ff) hevder at det premodifiserende leddet i norsk uansett lesing er et ikke-nominalt, modifiserende ledd, og at agentive nominaler alltid deler ut nullargument til agensrollen. Den språklige realiseringen av det premodifiserende leddet er da et pragmatisk spørsmål. Mens verbet "løpe" alltid vil tildele sitt grammatiske subjekt en agentiv rolle (*Per løper*), trenger ikke det korresponderende nominalet "løping" å gjøre det samme realisert som en premodifiserende possessiv (*Pers løping*).

Skillet mellom prosess og produktlesingen av nominaler tolket som to typer hendelsesdenotasjon synes i alle fall å ha noe med grunnleggende forskjeller mellom verbs og nominalers semantikk å gjøre. Verb typisk beskriver ulike situasjonstyper fra ulike aspektuelle og temporale vinkler, mens nominaler typisk refererer til korrespon-

derende typer. Grunnleggende elementer ved beskrivelse av situasjoner er faktorer som tid, rom og kausalitet (Widell 1996:139). Disse faktorene defokuseres og nedtones ved referanse:

12. *Arkitekten tegner plattformen.*
13. *Arkitektens tegning av plattformen.*

I 12 er kun tolkningen som den komplekse hendelsen tilgjengelig. Derfor er også agensleddet obligatorisk grammatikalisert gjennom det grammatiske subjektet. I 13 vil det korresponderende nominalet referere til en situasjon. Nominaler refererer typisk til tellbare statiske enheter. Situasjoner, inkludert hendelser, kan også konseptualiseres som tellbare statiske enheter ved nominalisering, jf at enkle hendelsesdenoterende nominaler kan pluraliseres, i motsetning til de komplekse:

14. *Arkitektens tallrike tegninger av plattformen.*

Pluraliseringen i (14) disambiguerer nominalet *tegnings* som entydig enkel hendelsesdenoterende.

Men verbalnominaler har også den egenskapen at de kan referere til både dynamiske situasjonstyper (*Pers løping*) og statiske (*Pers søvnn*), og gjennom referansefunksjonen gi dem et mer statisk preg. Muligheten for nullrealisering av premodifiserende possessivledd i agentiv funksjon kan sees på som en funksjon av nominalers statiske referanseegenskap. Dette er typiske trekk ved visse typer fagtekster (jf Andersen 1989/90).

Pustejowsky (1991, 1993) illustrerer hvordan det logiske potensial for semantiske roller også er til stede ved ikke-hendelsesdenoterende nominaler som *roman*. Alle nominaler, uansett type er omgitt av et system av relasjoner som gir en karakteristikk av nominalets semantikk. Dette systemet kaller han kvaliastrukturer. De likner semantiske roller, men er ikke statisk bundet til et leksem. Pustejowsky opererer blant annet med en telisk rolle, som angir referentens hensikt og funksjon, og en agentiv rolle, som angir referentens opprinnelse og tilkomst. Leksemer kan arve kvaliastrukturroller fra hyperonyme leksemer. Leksemet *roman* har for eksempel arvet sin teliske og agentive rolle fra hyperonymet *litteratur*. En anvendelse av Pustejovskys kvaliastrukturteori ville utvide rolleanalysen til å gjelde hele leksikon i et språk, og kunne bli meget omfattende. For begrepshierarkier innen domenebegrensede fagområder vil et slikt perspektiv derimot kunne virke mer lovende, især innen domener hvor det er mulig å etablere terminologiske typehierarkier (Andersen 1992).

3. Argumentstruktur og subspråk

Domenebegrensede fagområder kalles også subspråk (Sager et al. 1980) eller fagspråk (Hoffmann 1979). Siden de verbalnominalene som skal omtales, hører til en del av et domenebegrenset område, vil relevante aspekter ved fagspråksbegrepet bli omtalt nærmere.

Som mange fagspråksforskere har vært inne på er det typisk for fagspråk at et gitt subspråk av et allmennspråk representerer en begrenset delmengde av allmennspråkets morfosyntaktiske repertoar (jf Andersen 1989/90). Til gjengjeld benyttes denne delmengden av reper-toaret oftere enn i allmennspråket. Klassisk studerte fenomener er passiv, nominalisering, upersonlige konstruksjoner og setningers kompleksitet. Det er naturlig å søke etter funksjonalistiske forklaringer på dette, som Pragskolens tekstlingvistikk har gjort i stor utstrekning i flere år.

På grunn av det begrensede utvalget av morfosyntaktisk repertoar er det mest fremtredende inntrykk man får av tekniske tekster, for eksempel innen norsk petroleumsindustri til havs, deres statiske, monotone karakter. Både passivbruk, nominalisering og upersonlige konstruksjoner understreker dette. Et fellestrekk for disse konstruksjonstypene er at de gir muligheter for undertrykking av agensledd, noe som ikke bare gir et statisk inntrykk, men samtidig et inntrykk av "leseravstand" og objektivitet. Dette er et gjennomgående trekk ved beskrivende tekniske tekster, som den norske terminologiske databasen NOT henter en stor del av sitt grunnlagsmateriale fra. Fravær av agensrollen ved passiver og nominaliseringer er en del av den samme tendensen (Andersen 1998).

Databasen NOT er Norges største, og skal bygges ut fra å være en terminologisk base i et standard. termpostformat til å bli en relasjonell database (Meurer 1997). I tillegg skal verbalnominaliserte termer med tilhørende argumentstruktur innlemmes i en nasjonal elektronisk ordbok, som utvikles ved NTNU/Universitetet i Trondheim i det nasjonale satsingsprosjektet NorKompLeks (Norsk komputasjonelt leksikon).

4. Argumentstruktur og tilpasninger

Delmengdehypotesen om subspråk/fagspråk reiser en del viktige spørsmål om forholdet mellom grammatikk og leksikon, både i en allmennspråklig og i en fagspråklig ramme. Spørsmålet diskuteres av Pedersen (1995). Ifølge Pedersen må grammatikk og leksikon balanse slik at de nødvendige og relevante opplysningene nedfelles i en subspråksgrammatikk (kalt *differensgrammatikk*). Denne grammatikken vil ha store overlappingsområder med språkets allmenngram-

matikk, men i tillegg vektlegge og fremheve særtrekk ved et subspråk, for eksempel tellbarhet ved visse typer substantiver (op.cit.:130f).

En differensgrammatisk innfallsvinkel slik Pedersen foreslår, synes å være fruktbar i denne sammenhengen. Ved inkorporering av NOT-basen i NorKompLeks-materialet vil argumentstrukturer ved terminologiske verbalnominaliseringer naturlig sortere under en differensgrammatikkens beskrivelsesområde. Mer generelt blir da følgende opp-gaver sentrale:

- tilpassing av differensgrammatikk og teksttypologi/genre*
- tilpassing av differensgrammatikk og terminologisk leksikon*
- tilpassing av differensgrammatikk og allmenngrammatikk*
- tilpassing av terminologisk database til terminologisk leksikon*
- tilpassing av terminologisk leksikon og allmennleksikon*

a) og b) henger nøye sammen, særlig ved oljetermene i NOT, som har en bestemt tekstmengde som basis, og det er naturlig å anta at data fra a) er en viktig forutsetning for b). Av de verbalnominaliseringene som hittil er analysert i NOT-basen innen delområdene boring og prosess (vedlikehold er ikke ferdig analysert i skrivende stund), har samtlige urealisert premodifiserende agensledd (*(*arkitektens) tegning av plattformen*). Det totale antall norske hovedtermer innen boring er 1951. Av disse er 168 registrert som verbalnominaliseringer med argumentstruktur. Innen prosess er totalantallet norske termer 2453, og av disse er 300 registrert som verbalnominaliseringer med argumentstruktur.

Dette tyder på at det premodifiserende leddet ikke er fakultativt, som i allmennspråket, men er systematisk undertrykt i denne typen tekster. Lødrups pragmatiske analyse synes å kunne harmoniseres med dette: Genremessig og teksttypologisk variasjon kan sees på som en pragmatisk/stilistisk variabel. I visse gener, som i typiske allmennspråklige gener, er det premodifiserende leddet i verbalnominaler fakultative, og realiseres hvis det er fare for misforståelser, mens andre gener, som visse fagspråklige domener, nedfelt i visse teksttyper (som beskrivende tekniske tekster) vil agensleddet obligatorisk og dermed systematisk undertrykkes.

I b) vil derfor en differensgrammatikk inneholde informasjon om at verbalnominaler innenfor domenet ikke har foranstilte agentivledd. I det fagspråklige leksikon vil nominalenes argumentstruktur kun angis for postmodifiserende ledd, som *tegning av plattformen*. Hvis tilpassing mellom en differensgrammatikk og det terminologiske leksikon skal oppnås, må NOT-basen bygges ut med morfologisk informasjon, dvs. angivelse av ordklasse, samt bøyingsparadigmer for tall og kjønn, som i vanlige allmennspråklige ordbøker.

I NOT vil dette også medføre, som et aspekt av punkt d), at de tallrike komplekse termene (flerordstermer) må lemmatiseres, fordi det kun er den komplekse termens kjerneledd som skal være inngangsterm og dermed være bærer av de grammatiske bøyingsparadigmene. Lemmatisering i NOT krever ikke bare lingvistisk, men også faglig kompetanse innen det aktuelle fagområdet. Ved lemmatisering må man gå tilbake til de opprinnelige engelske termene og foreta en IC-analyse av dem. Slike analyser i NOT-basen har vist at mange komplekse termer er strukturelt tvetydige. La oss betrakte noen eksempler:

Den engelske termen *fire damper status lamp* vil for eksempel ha to mulige IC-analyser som resulterer i to mulige norske termkandidater. Disse alternativene har konsekvenser for lemmatiseringen: I den ene analysen vil kjernen være *lamp* og *fire damper status* vil være premodifiserende ledd. Den norske oversettelsen vil da være *lampe for brannspjeldstatus*, og *lampe* vil bli det lemmatiserte leddet. I den andre analysen vil kjernen være *status lamp*, og *fire damper* vil være modifierende ledd. Den norske oversettelsen vil i så fall være *statuslampe for brannspjeld*, og det lemmatiserte leddet vil være *statuslampe*.

Det er utelukkende en innsikt i de faglig/pragmatiske forholdene og konteksten som kan sette oss i stand til å foreta den korrekte lemmatiseringen i disse tilfellene. I NOT vil dette gjelde for komplekse termer som *helicopter fuel tank* (*tank for helikopterbrennstoff* eller *brennstofftank for helikopter?*) og *material classification code* (*kode for materiellklassifisering* eller *klassifiseringskode for materiell?*).

Korrekt lemmatisering ved nominaliserte komplekse termer er også viktig i en tilpassing mellom differensgrammatikk og terminologisk leksikon (punkt b): Også her er det kjerneordet, verbalnominalet, som skal være bærer av det grammatiske bøyingsparadigmet, og det er jo med utgangspunkt i verbalnominalet at argumentstrukturangivelsene skal angis. Hovedtypene kan angis i en differensgrammatikk, mens leksikonoppslaget enten kan referere til hovedtypene ved hjelp av koder (som i NorKompLeks), eller de angis i full form i leksikon.

Siden det ikke foreligger en utbygd differensgrammatikk for domenet, er den sistnevnte løsningen valgt i NOT. De to andre spørsmålene som er forbundet med punkt b), forholdet mellom produkt og prosesslesingen av verbalnominalene, og forholdet mellom komplekse verbalnominaltermer og sammensatte termer, blir drøftet nærmere nedenfor.

Når det gjelder c): tilpassingen mellom differensgrammatikk og allmenngrammatikk, er forholdet slik at det foreligger en norsk referansogrammatikk (Faarlund et al 1997) som det er naturlig å forholde seg til. Referansogrammatikken (op.cit.:717) behandler semantiske roller i forbindelse med verb, og setter dem opp i et klassisk tilgjengelighetshierarki:

1. agens, 2. recipiens, 3. patiens.

Patiens deles i affisert (*reparasjon av kompressoren*) vs ikke-affisert (*lesing av dokumenter*). Ideelt sett ville man da forvente at det ville være et tilsvarende hierarki i NorKomLeks med en eventuell videre spesifisering av roller etter deskriptivt behov. Vi ser imidlertid at patiensrollen i referansogrammatikken er byttet ut med andre, og at delvis andre etiketter er brukt:

1. Agent, 2. Theme, 3. Goal, 4. Benefaktiv, 5. Recipient, 6. Malefaktiv, 7. Experiencer, 8. Instrument, 9. Lokativ.

Etter hvert som spesifikasjonsbehovet i analysene har økt, har behovet for ytterligere utvidelse av antallet roller meldt seg, for eksempel *resip*, *norole* ("no role") og *unspec*. Vi synes med andre ord å stå overfor det dilemmaet som ble illustrert gjennom Dowty-sitatet i første del av denne artikkelen.

I en tilpassing mellom Norsk referansogrammatikk og NorKompLeks har vi valgt en mellomløsning hvor vi i utgangspunktet ønsker å begrense antallet tematiske roller mest mulig. Vi har også søkt å ta høyde for at rolleetikettene og deres definisjoner harmonerer med NorKompLeksrollene, slik at vi tillater oss å slå sammen roller og gi dem nye navn, men på en slik måte at de lett kan ekspanderes hvis ytterligere spesifikasjoner er nødvendige. I den følgende oversikten (Fjellheim 1998) er rollen *måte* innført selv om denne ikke er angitt i NorKompLeks og Norsk referansogrammatikk. Etikettene i klammeparentesene er identisk med de kodene vi bruker i NOT-basen ved registreringen:

1. [mål] MÅL:

Endepunkt for en bevegelse/handling. Deltaker/objekt som blir direkte påvirket av handlingen. Den/det som blir karakterisert av setningen, som er på et sted, skifter oppholdssted, eller blir til ved prosessen. Som i Norsk referansogrammatikk opererer vi med to undergrupper:

[mål] AFFISERT:

Objekt som blir direkte påvirket av verbalhandlingen (*nedpumping av utstyr*).

[resultat] RESULTAT (EFFISERT)

Objekt som blir til ved verbalhandlingen (*fremstilling av damp*).

2. [instr] INSTRUMENT

Den umiddelbare fysiske årsaken til en hendelse. Instrumenter, hjelpemidler som fungerer som formidler mellom agens og mål (inkluderer kroppsdel, mot styrke osv.) (*boring med stiv streng*).

3. [sted] STED

Sted der handling finner sted, ender eller kommer fra. Vi opererer med tre undergrupper: påsted PÅSTED (*boring på dyp, kjernetaking i sidevegg*), tilsted TILSTED (*fordeling av ferskvann til servicesystemet*) og frasted FRASTED (*drenering av vann fra havbunn*).

4. [måte] MÅTE

Angir en nærmere beskrivelse av handlingen. Inkluderer omstendigheter, som hvordan, når, osv. (*måling under produksjon*).

Måterollen er svært vanskelig å avgrense, og den glir lett over i andre mer temporale avskygninger, men vi ønsker å operere med den foreløpig.

5. Angivelse av argumentstrukturer ved verbalnominaler i NOT

Denne delen av artikkelen inneholder en oversiktsbeskrivelse av måten argumentstrukturene ved verbalnominalene er kodet på i NOT. Utgangspunktet er NorKompLeks sitt kodeverk. Systemet er programmert i LISP, som har en deklarativ syntaks som passer godt for formålet. Det grunnleggende i den deklarative syntaksen er at kodene inngår i et hierarki av attributtverdistrukturer (også kalt trekkstrukturer). Det legges vekt på å skille skarpt mellom tre analysenivåer: syntaktisk funksjon (f), semantisk rolle (r), syntaktisk form (c). Ved verbale strukturer vil f-rollene være funksjonskategorier som subjekt og objekt. I komplekse NP-ers interne argumentstruktur vil f-rollene være funksjoner som *premodifisert* og *postmodifisert* med hensyn til plassering relativt til det nominale kjerneleddet. De semantiske rollene (s-roller) er identiske med de som er angitt ovenfor. De syntaktiske formrollene (c-roller) vil enten være NP-er eller PP-er. I sistnevnte tilfelle angis kjerneordet (preposisjonen) eksplisitt:

```
operatørens nedpumping av utstyr i brønn
arg_str: [
arg1 :[f:premod, r:ag, c: np [cat:gen]]
arg2 :[f:postmod1, r:mål, c:[cat:pp, kjerne : "av"]]
arg3 :[f:postmod2, r:sted, c:[cat:pp, kjerne : "i"]]
```

Fordi det premodifiserende leddet alltid er tomt, og agens alltid er undertrykt, vil alle argumentene være etterstilte, og koding av f-rolle vil være redundant. Videre velger vi å holde oss til norske forkortinger, dvs. *k* for *kategori* i stedet for *c* for *category*. Kodingen kan dermed reduseres til følgende:

nedpumping av utstyr i brønn

```
arg_str: [
arg1 :[r:mål, k:[kat:pp, kjerne : "av"]]
arg2 :[r:sted, k:[kat:pp, kjerne : "i"]]
```

Det vil fremgå av argumentstrukturenes nummerering hvilken posisjon de har.

Vi kommer likevel ikke unna spørsmålet om hvor mange roller vi trenger, og hvilke etiketter vi i siste instans skal bruke. Dette vil uansett være et spørsmål om hva det er fruktbart å operere med i et gitt prosjekt. Tilpassingen til andre prosjekter er det avgjørende punktet her.

Når det gjelder nominalenes rolletildelingsrammer, får vi de samme problemene her som ellers. Ved et verb som *plassere* vet man at det er vanlig å analysere det som et 3-plasspredikat hvor det agentive subjektet, det direkte objekt med målrolle og preposisjonsleddet med tilstedsrolle alle faller inn under rolletildelingsrammen. Da blir *plassering* en 2-plassnominalisering hvor preposisjonsleddet med målrolle og preposisjonsleddet med tilstedsrolle begge hører til rolletildelingsrammen, som i *plassering av drivrør i drivrørshylse*. *Nedpumping* som i *nedpumping av utstyr i brønn* vil få tilsvarende analyse.

I tilknytning til punkt b), tilpassingen mellom differensgrammatikk og terminologisk leksikon, er det dessuten to forhold som må beskrives nærmere: forholdet mellom produkt- og prosesslesingen av verbalnominalene, og forholdet mellom komplekse verbalnominaler og sammensatte termer. Disse punktene har også å gjøre med punkt d), tilpassing av terminologisk database og terminologisk leksikon.

La oss ta et NOT-eksempel:

15. utkopling ved nødavstenging

Ved prosesslesing har (15) ett argument realisert som PP med kjerne preposisjonen *ved*. Den tilhørende semantiske rolle vil være måte. Dermed er dette et eksempel på en kompleks term, og *utkopling* vil bli lemmatisert i det terminologiske leksikonet. Ved produktlesing av (15) vil *utkopling* referere til det instrumentet hvor utkopling foretas. I så fall blir *ved nødavstenging* tolket som et valensuavhengig ledd som

uttrykker sin rolle uavhengig av *utkopling* (for eksempel en påstedsrolle, dvs. at utkopligen befinner seg i nærheten av et instrument som utløser nødavstenging). Ved produktlesingen vil dermed (15) ikke være en term. Den naturlige terminologiske analysen er da å registrere *utkopling* som en term og *nødavstenging* som en annen term.

Denne hypotesen aktualiserer altså spørsmålet om argumentstrukturanalysens relevans for termidentifikasjon. Jeg går ikke ytterligere inn på dette spørsmålet her, men anser det for å være et problem som det er verd å forfølge.

Forskjellen mellom prosess- og produktlesingen av nominalene fører til at de i terminologisk analyse må realiseres som to ulike termer, fordi de representerer to ulike begreper. Spørsmålet er også om de må representeres som to ulike leksikalske oppslag eller kun ett. Pustejovsky (1991:432) analyserer denne tvetydigheten ved nominaler som en gjennomgående, systematisk egenskap ved verbale nominaler på lik linje med tellbarhet vs ikke-tellbarhetsdenotering og ulike metonymidenotasjoner. Denne systematikken tyder på at beskrivelsesbyrden her snarere må legges til en grammatikk, og at man kun trenger ett leksikalsk oppslag, hvor produkt- og prosesslesing fremgår av eksempler i leksikonartikkelen (jf skillet mellom grammatikk og leksikon i modell 1 ovenfor). Ved terminologisk analyse, derimot, må man for hvert enkelt tilfelle identifisere om prosesslesing foreligger eller ikke.

Siden dette synes å være et kontekstsensitivt fenomen, må det i så fall være et problem for en terminologisk teori som nedfeller enetydighetsprinsippet og termers kontekstfrihet som grunnleggende prinsipper. Det er jo nettopp konteksten (inkludert den faglige) som avgjør hvilken lesning vi får. En dobbeltoppføring for hver lesning synes kontraintuitivt når den tydeligvis er en grunnleggende logisk egenskap ved nominalet selv. Det synes derfor som om skillet snarere er på et pragmatisk tekstnivå enn på termnivå. Forskjeller mellom tekstnivå og termnivå skaper ofte identifikasjonsproblemer ved termekserpering. Det gjelder også identifikasjonen av komplekse hendelsesdenoterende nominaler som termer, som vi har sett. Særlig gjelder dette i de tilfellene hvor åpenbart valensbundne ledd ikke realiseres i tekst. (15) er et eksempel på dette. *Utkopling* har et korresponderende verb *kople ut*, som åpenbart krever et direkte objekt. Ved nominalisering realiseres tilsvarende ledd ved preposisjon *av + styrt ledd*. Grunnene til at et valensbundet ledd ikke realiseres tekstlig, er pragmatiske forhold, som for eksempel relasjonen mellom gitt og ny informasjon i tekster. I disse tilfellene er det derfor åpenbart at skillet mellom termnivå og tekstnivå resulterer i angivelse av urealiserte argumenter på termnivå. Argumentstrukturen til (15) vil da angi urealiserte argumenter som i parentes: *utkopling (av x) ved nødavstenging, skjøting (av x) i rørkoplingshylse, etc.*

Ifølge Grimshaw (1990:68f) kan argumentstrukturanalyse også anvendes på sammensatte ord hvor kjerneleddet er et komplekst hendelsesdenoterende verbalnominal. Disse kalles **syntetiske sammensetninger**, til forskjell fra **rotsammensetninger**. De førstnevnte har argumentstruktur til forskjell fra de sistnevnte. På engelsk korresponderer dette skillet mellom -ing-nominaliseringer og nominaliseringer dannet av verbets rot. Skillet synes til en viss grad å stemme også for norsk. Et viktig poeng er at ved syntetiske sammensetninger kan ikke agens opptre som argument, til forskjell fra tilsynelatende agentvaktige argumenter ved rotsammensetninger; noe som vil være i tråd med føringene på de komplekse verbalnominalene hvor agens systematisk undertrykkes:

16. *biestikk, hundebitt, hjertebank*
17. *?biestikking, ?hundebiting, ?hjertebanking*

Forholdet mellom de to typene nominalisering er mer usikker på norsk, men forholdet mellom syntetiske sammensetninger og verbalnominaliseringer med prosesslesing og tilhørende argumentstruktur synes å stå i et systematisk forhold til hverandre også på norsk:

18. *avstenging av brønnramme*
19. *brønnrammeavstenging*
20. *dosering av biocid*
21. *biociddosering*

(18) og (19) er ekvivalente termer, og (20) og (21) er ekvivalente termer. Dette systematiske forholdet gjennomsyrrer de registrerte -ing-nominaliseringene med prosesslesing i NOT. La oss kalle de komplekse termene (flerordstermene) **utpakkede termer** (dvs. (18) og (20), og sammensetningstermene **innpakkede termer** (dvs. (19) og (21)). Forholdet mellom innpakkede og utpakkede termer har vært studert i NOT-materialet på et tidligere tidspunkt (Myking 1988). I terminologisk sammenheng synes det å være visse føringer på forholdet mellom de to termvariantene. Dette har med begrensninger på mulig ordlengde (og dermed hvor mye som kan pakkes inn) i norsk å gjøre:

22. *dosering av flokkuleringsmiddel*
23. *?flokkuleringsmiddeldosering*
24. *fordeling av ferskvann til servicesystemet*
25. **servicesystemferskvannfordeling*

Likevel synes det klart at det er et regelmessig derivasjonelt forhold mellom utpakket og innpakket term. En viktig forskjell er at de inn-

pakkede termene er **underdeterminerte** i forhold til de utpakkede, dvs. at de utpakkede inneholder mer lingvistisk informasjon (signalisert særlig gjennom preposisjonene), som "forflates" ved innpakkingsprosessen. Derfor representeres argumentstrukturene for utpakket form også ved sammensetninger. Hvis en term som *brennerstyring* er registrert, vil den "underliggende" utpakkede termen *styring av brenner* danne utgangspunktet for argumenstrukturbeskrivelsen:

26. *brennerstyring*
arg1: [r: mål, k:[kat: pp, kjerne "av"]]

På denne måten kan vi automatisk pakke inn termer, og studere deres struktur i lys av de føringene på ord som ble nevnt ovenfor.

6. Konklusjon

I denne artikkelen har jeg prøvd å vise hvordan argumentstrukturteori i ulike former kan bidra til å gi nye perspektiver på terminologisk og leksikografisk tenking, og hvordan man stadig må vurdere forholdet mellom grammatikk og leksikon (inkludert fagleksikon/terminologi). Disse forholdene er det spesielt viktig å vurdere i forbindelse med subspråks leksika for å tilrettelegge dem, ikke bare for profesjonelle brukere innen det aktuelle fagområdet, men også for terminologi og fagspråksforskere til ulike datalingvistiske formål. Det bør være mulig å forene disse to behovene i en leksikalsk database med relasjonell oppbygging, slik at flere brukergrensesnitt kan bygges inn i den.

7. Litteratur

- Andersen, Øivin 1989/90: Fagspråk som forskningsdisiplin. I: *Hermes. Journal of Linguistics* nr 3-1989, 73-99, nr 4-1990, 85-101. Handelshøyskolen i Aarhus.
- Andersen, Øivin 1992: Begrepshierarkiers logiske og empiriske egenskaper. I: *Terminologi, Edb og Vidensteknik*. Nordterm Proceedings. Varde, 33-54.
- Andersen, Øivin 1997: Not-prosjektet: Norsk terminologisk database – Struktur og applikasjoner. I: Hjulstad, Haavard (red.): *Nordterm '97*. Oslo: Nordterm/RTT. (= Nordterm 9.), 61-66.
- Andersen, Øivin 1998: Fagtekst og agentivitet. I: Laurén, Christer/Nordman, Marianne (red.): *Nation och individ i Fackspråk. Forskning i Norden*. Vasa, 12-27.

- Andersen, Øivin/Johan Myking 1997: Norsk terminologisk database – kontekst og validering. Prosjektbeskriving, Norsk termbank.
- Benson, Michael 1989: The Structure of the Collocational Dictionary. I: *International Journal of Lexicography*. 2.1. Oxford University Press. Oxford.
- Chomsky, Noam 1970: Remarks on Nominalization. I: Jacobs, R.A./Rosenbaum, P.S. (red.): *Readings in English Transformational Grammar*. Ginn and Company. Waltham, 184-221.
- Dowty, David 1989: On the Semantic Content of the Notion of "Thematic role". I: Chierchia, G./B.H. Partee/ R. Turner (red.): *Properties, Types and Meaning, II*. Dordrecht:Kluwer Academic Publishers, 69-129.
- Faarlund, Jan Terje/Svein Lie/Kjell Ivar Vannebo 1997: *Norsk Referansegrammatikk*. Universitetsforlaget. Oslo.
- Fjellheim, Hilde 1998: Argumentstrukturer for nominaliseringer i NOT. Arbeidsnotat. Norsk termbank.
- Grimshaw, Jane 1990: *Argument Structure*. Linguistic Inquiry. Monograph 18. The MIT Press, Cambridge, Mass.
- Heltøft, Lars 1996: Grammatikalisering af semantiske roller i dansk. I: Rasmussen, Lone Schack (red.): *Semantiske roller*. Odense Working Papers in Language and Communication. 10. Odense, 43-65.
- Hoffmann, Lothar 1979: Towards a Theory of LSP. I: *Fachsprache* 1.2, 12-17.
- Lødrup, Helge 1989: *Norske Hypotagmer. En LFG-beskrivelse av ikke-verbale hypotagmer*. Novus forlag. Oslo-studier i språkvitenskap 4.
- Meurer, Paul 1997: Et verktøy for terminologiske databaser. I: Hjulstad, Håvard (red.): *Nordterm '97*. Oslo: Nordterm/RTT. (= Nordterm 9.), 66-77.
- Myking, Johan 1988: Complex Noun Phrases as a Problem of Terminological Practice. I: Laurén, Christer/Marianne Nordman (red.): *Special Language. From Humans Thinking to Thinking Machines*. Multilingual Matters. LTD. Clevedon. Philadelphia, 265-275.
- Pedersen, Jette 1995: Grammatiske opplysninger i tekniske fagordbøker. I: *LexicoNordica*. 2, 121-143.
- Pitz, Anneliese 1994: *Nominal Signs in German*. Dr. art.-avhandling. Lingvistisk institutt, Det historisk-filosofiske fakultet, Universitetet i Trondheim.
- Pustejovsky, James 1991: The Generative Lexicon. I: *Computational Linguistics*. 17.4, 409-441.
- Pustejovsky, James/Branimir Boguraev 1996a: *Lexical Semantics. The Problem of Polysemy*. Clarendon Press. Oxford.

- Pustejovsky, James/Pierette Bouillon 1996b: Apectual Coercion and Logical Polysemy. I: Pustejovsky, James/Branimir Boguraev (red.) 1996a, 133–163.
- Rappaport, Malka 1983: On the Nature of Derived Nominals. I: Levin, L./M. Rappaport/A. Zaenen (red.): *Papers in Lexical-Functional Grammar*. Indiana University Linguistics Club. Indiana, 113–142.
- Sager, Juan Carlos/D. Dungworth/P.F. McDonald 1980: *English Special languages*. Wiesbaden.
- Widell, Peter 1996: Apektuelle verbalklasser og semantiske roller – Den dobbelte aspektkalkule. I: Rasmussen, Lone Schack (red.): *Semantiske roller*. Odense Working Papers in Language and Communication. 10. Odense, 135–169.