

RETNINGSLINJER FOR SYNTAKTISK ANNOTASJON I LIA

Andre Kåsen, Kristin Hagen, Lilja Øvreliid, Signe Laake, Håvard Østli

April 2019

Annotasjonen i LIA skal følge NDT og «Retningslinjer for morfologisk og syntaktisk annotasjon i Norsk dependenstrebanks» (Kari Kinn, Per Erik Solberg og Pål Kristian Eriksen, 2013) så langt det er mulig.

Denne veiledningen fokuserer på den syntaktiske annotasjonen av talemålstrekk hentet fra transkripsjonene i LIA. Talemål skiller seg klart fra skriftspråk og har mange trekk som ikke finnes i skriftspråk. Veiledningen er ment som et supplement til retningslinjene for NDT. Framstillingen tar utgangspunkt i The Universal Dependencies Treebank of Spoken Slovenian (Dobrovolje & Nivre, 2016).

Innhold

1	Oppdeling i segmenter/setninger	3
1.1	Flere finitte verb i et segment:.....	3
1.2	KOORD-ELL	4
1.3	«Tomme» segmenter	5
2	Extralinguistic Tokens	5
2.1	Setningsskilletegn	5
2.2	Setningsinterne skilletegn	5
2.2.1	Pauser	6
2.2.2	Anførselstegn.....	7
2.3	Avbrutte ytringer	8
3	Disfluencies.....	9
3.1	Repairs	9
3.2	Restarts	10
3.3	Ufullstendige ordstrenger	11

4	Ufullstendige ytringer.....	13
4.1	Ufullstendige ytringer i begynnelsen av et segment.....	13
4.2	Ufullstendige ytringer i slutten av et segment.....	15
4.3	Ytringer uten finitt verbal.....	16
4.3.1	INTERJ:.....	16
4.3.2	FRAG:.....	20
5	Discourse elements	21
5.1	Interjeksjoner	21
5.2	Discourse fillers	23
6	Parentetiske uttrykk.....	24
6.1	Direkte tale	24
6.2	Andre typer utsagnsverb.....	24
6.3	Metakommentarer	26
7	Apposisjoner.....	27
7.1	Subjektsdublering.....	27
7.2	APP når et foranstilt ledts innhold gjentas med så eller <i>da</i> (4.2.4. i NDT)	27
8	Formelle subjekter.....	28
9	FINV og INFV.....	29
10	Ellipser	30
11	Setningsknuter og preposisjonsstranding	30
12	Andre frekvente talemålstrekk.....	32
12.1	Prepropriell artikkel:.....	32
12.2	Konjunksjon på slutten av et segment	32
12.3	Da på slutten av segment	33
12.4	Oppramsing, opplesning av lister	34
13	Frekvente «faste» uttrykk.....	35

1 Oppdeling i segmenter/setninger

Vi følger transkripsjonen og lager ikke nye setningsskiller.

Eksempel

90 - ja .

91 - eg har kanskje e vore nokså sterkt knytt til skolelaget på Ve .

92 - e på ungdomsskolen der # som er Norges kristelige student- og skoleungdomslag .

93 - da e da eg gjekk der på skolen så # på ungdomsskolen så begynte eg jo der som medlem og .

94 - så blei jo formann den siste tida og så nå # her i fjor så var dei # forlegne for # leiarar så da begynte eg igjen .

95 - saman med mannen min # så vi har det nå .

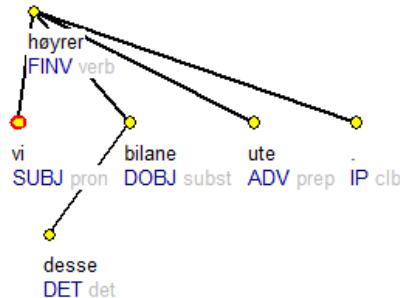
96 - som vi driv med det er jo # ei gong i veka så er det jo styremøte som ein må vere med på ## så det .

97 - er det andre slags e # ting du er engasjert i ?

98 - ikkje nå e # for augeblikket altså .

Punktumet på slutten av hvert segment er satt inn automatisk av Oslo-Bergentaggeren som er brukt til å tagge treningskorpuset morfologisk. Vi beholder punktumene i den syntaktiske annotasjonen siden vi planlegger å bruke (deler av) NDT som treningsmaterialet for LIA.

Dette punktumet skal henge på øverste node i treeet og ha funksjonen IP.



(1) vi høyrer desse bilane ute.

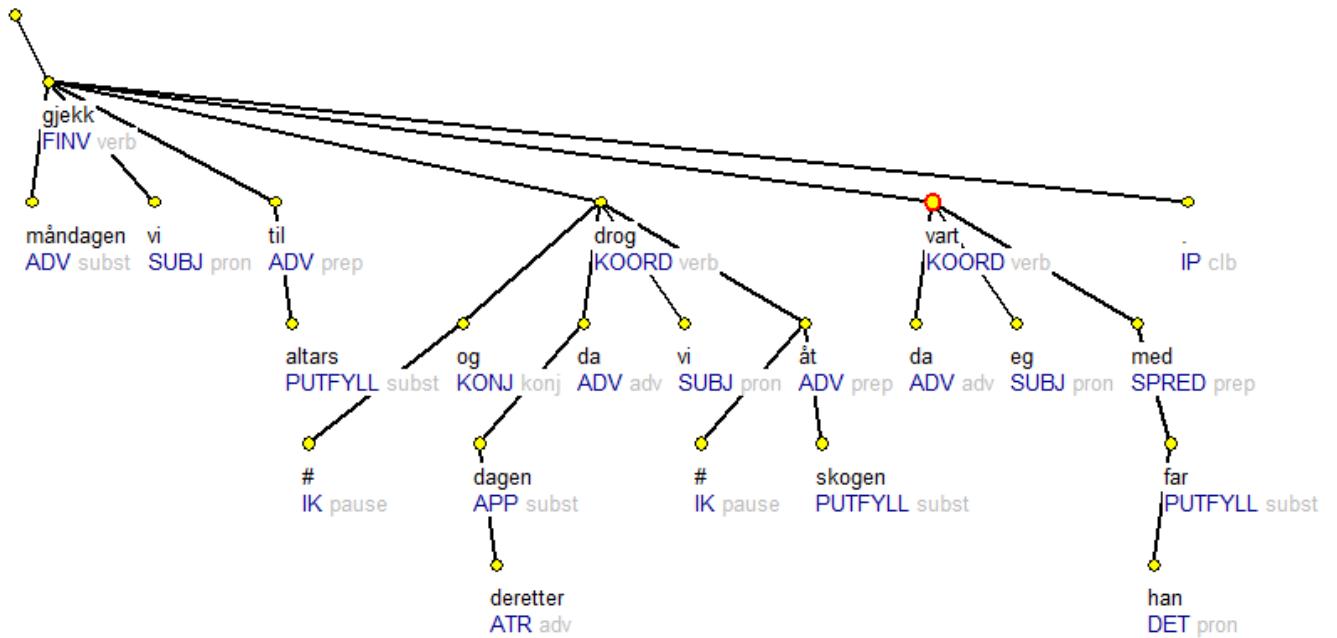
1.1 Flere finitte verb i et segment:

Fordi det ofte er mer enn et finitt verb i et segment, bruker vi funksjonen KOORD for å koordinere de finitte verbene. Det første finitte verbet i den første «helsetningen» i segmentet får funksjonen FINV, mens de andre etterfølgende finitte verbene får funksjonen KOORD. Hvis de etterfølgende finitte verbene er koordinert på samme nivå (altså det er ikke noen underordning), skal alle de etterfølgende verbene henge på FINV.

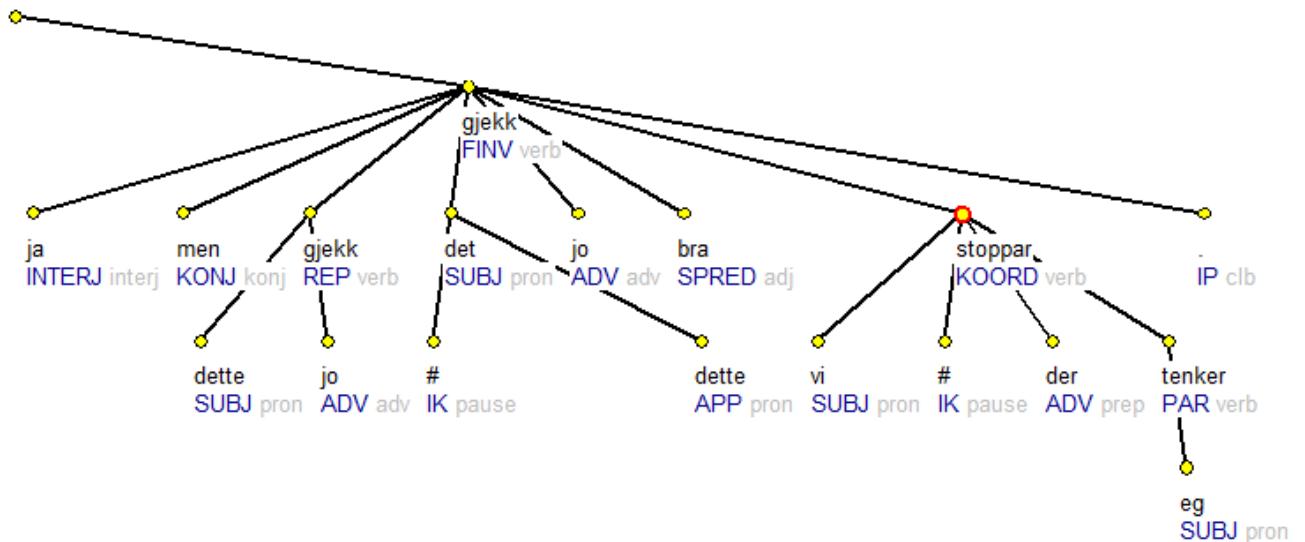
Det trenger ikke å være en innholdsmessig sammenheng mellom de koordinerte setningene.

Her skiller våre retningslinjer seg fra NDT-retningslinjene. Det er viktig at de to koordinerte leddene har samme funksjon i setningen. Det er ikke alltid at dette kravet blir oppfylt hos oss.

(2) «måndagen gjekk vi til altars # og dagen deretter da drog vi # åt skogen da vart eg med han far ..» (aamot_uio_01 ... 0000 37/423)¹



(3) «ja men dette gjekk jo # det gjekk jo bra dette vi # stopper der tenker eg ..»
(austevoll_uib_01 ... 0007 1/20)»



Utsagsverb skal ha funksjonen PAR. Se kapittel 6 for mer om dette.

1.2 KOORD-ELL

I NDT brukes funksjonen KOORD-ELL hvis det finitte eller infinitte verbet er utelatt ved koordinasjon:

(1) Han ble boende hos faren, storesøstera hos mora. (NDT s. 119)

¹ Vi refererer til de aktuelle segmentene i filmaterialet på denne måten, dvs. filnavn, filnummer, segment. De tre prikkene indikerer at vi har utelatt, i denne sammenhengen, unødvendig informasjon, for å gjøre det enklere å referere.

KOORD-ELL er ikke en funksjon vi bruker i LIA. Funksjonen er ikke tilgjengelig i tag-settet, og vi bruker KOORD i stedet.

1.3 «Tomme» segmenter

Noen segmenter er tomme og inneholder bare et punktum eller andre skilletegn. Treet vil da se for eksempel slik ut:



2 Extralinguistic Tokens

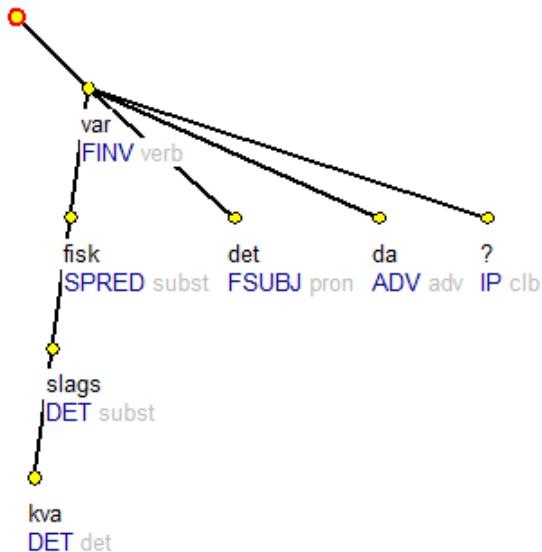
Hovedregel: Skilletegn behandles som NDT. Se 5.2.

2.1 Setningsskilletegn

Punktum er behandlet i kapittel 1, ellers finnes bare «?» i LIA-transkripsjonene.

Disse skal være dependent på rotnoden, og ha funksjonen IP.

- (4) Kva slags fisk var det da? (aamot_udio_01 ... 0000 287/423)



2.2 Setningsinterne skilletegn

I LIA finnes det to typer setningsinterne skilletegn: pausetegnet # og anførselstegnene «»

Setningsinterne skilletegn får funksjonen IK som i NDT.

Vi følger, i stor grad, NDT (s. 125-9) som har følgende prinsipper for setningsinterne skilletegn:

- a) Skilletegnet er dependent på hodet til en av de syntaktiske enhetene det skiller ut.

Om skilletegnet står på venstre eller høyre enhet, er avhengig av de ulike typene av skilletegn. Se 2.2.1 for pauser og 2.2.2 for anførselstegn.

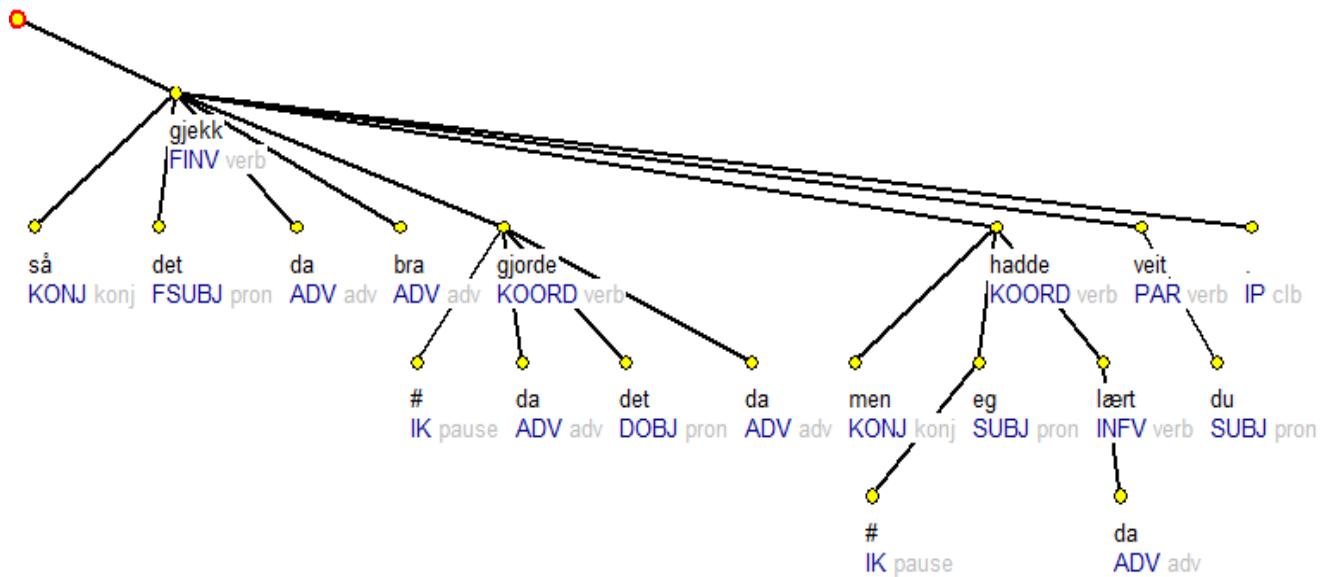
- b) I tilfeller der det er flere løsninger, skal skilletegnet stå så høyt som mulig så lenge projektivitet bevares. (Dobrovolje & Nivre, 2016: 1568). Ved ikke-projectivitet vil det være kryssende kanter. Det betyr at man må henge dependenten så høyt som mulig, men likevel så lavt at man unngår kryssende kanter.

2.2.1 Pauser

Pauser skal så langt det lar seg gjøre, være dependent på den syntaktiske enheten til høyre i strengen.

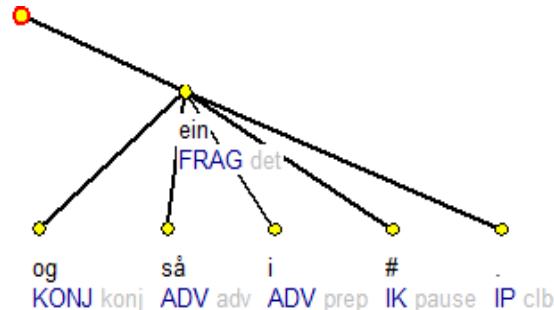
Med syntaktisk enhet mener vi ordet som står til høyre i strengen.

- (5) så det gjekk da bra # gjorde da det da men # eg hadde lært da veit du. (aamot_udio_01 ...
0000 22/423)



Hvis det ikke finnes noe ord til høyre i strengen som det er naturlig å henge tegnet på, skal det henges på rotnoden.

(6) og så ein i #. (aamot_udio_01 ... 0000 243/423)

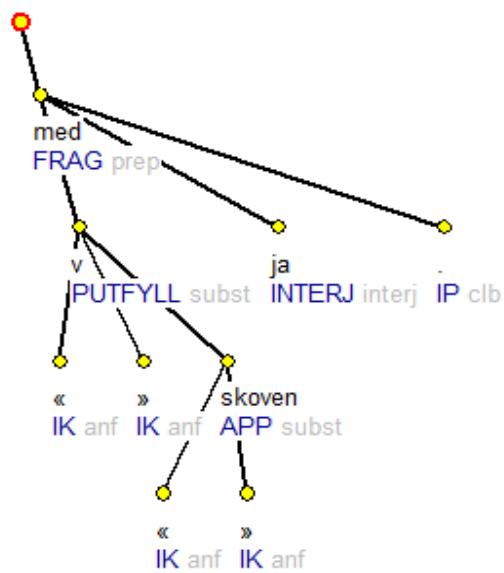


Vi behandler nølelyder likt som pauser. For mer om nølelyder, se 5.2.

2.2.2 Anførselstegn

Anførselstegn behandles litt annerledes. Det første anførelstegnet henges på ordet til høyre i strengen. Det siste anførelstegnet henges på ordet til venstre i strengen.

(7) med «v» «skoven» ja. (aamot_udio_01 ... 0000 228/423)



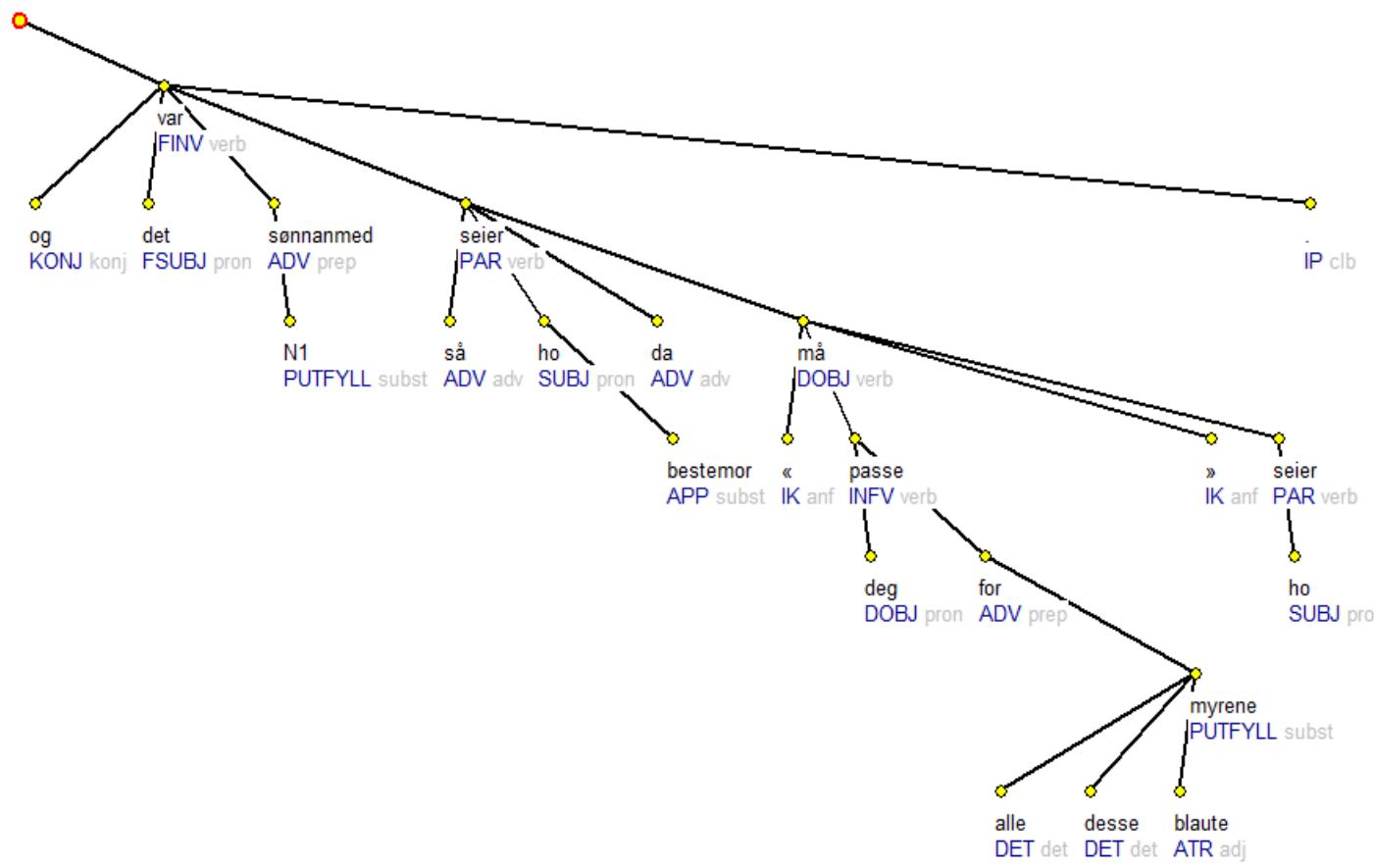
Når anførelstegnene avgrenser mer enn et ord, bruker vi samme regel som i NDT (5.2.4):

«Åpningsanførelstegn skal være dependent på hodet på det utskilte elementet til høyre, mens sluttanførelstegn skal stå på hodet til venstre. Sagt på en annen måte: de to korresponderende

tegnene skal være dependenter på samme hode.» Det betyr at begge anførselstegnene skal henge på det høyeste hodet i det de avgrenser.

I eksempel (8) avgrenser anførselstegnene denne delen av ytringen: *må passe deg for alle desse blaute myrene*. Det høyeste hodet er *må*, og begge anførselstegnene skal henge her.

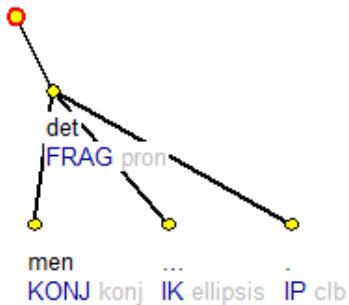
(8) og det var sønnanmed N1 så seier ho da bestemor «**må passe deg for alle desse blaute myrene»** seier ho . (aamot_uio_01 ... 0000 8/423)



2.3 Avbrutte ytringer

Dersom en ytring ikke blir skikkelig avsluttet, og en annen taler tar over, for eksempel ved avbrudd, skal den avbrutte ytringen avsluttes med tre punktum med mellomrom foran: ... (tre prikker) i transkripsjonen. Disse tre prikkene får funksjonen IK og er dependenter på rotnoden.

(9) men det (aamot_udio_01 ... 0000 375/423)



3 Disfluencies

3.1 Repairs

Eksempel: og e # dei var øq # dei følgde øq godt med alt # utanom

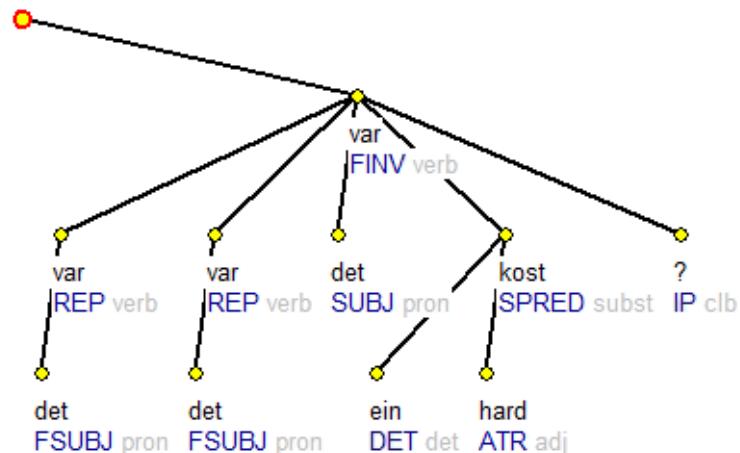
↓ ↓

Dobrovolje	Reparandum	repair
& Nivre:	dependent	
	på sin repair	
	(høyeste node)	

Repair kan være *repetition, substitution, reformulation*. Det er den innholdsmessige likheten mellom de to delene som bestemmer at den ene delen er en REP. Det betyr at en REP kan være en fullstendig syntaktisk enhet, for eksempel en setning med et finitt verb.

I LIA gir vi reparandum-leddet funksjonen REP og setter den på høyeste node i repair-delen.

(10) det var det var det var ein hard kost ? (aamot_uio_01 ... 0000 105/423)



Ikke bruk mye tid på å lage en syntaktisk struktur innad reparandum. Den syntaktiske strukturen kan gjerne være flat.

Hvis det ikke fremkommer klart at reparandum og repair er innholdslike, så er relasjonen mellom de to delene SLETT (se 3.2).

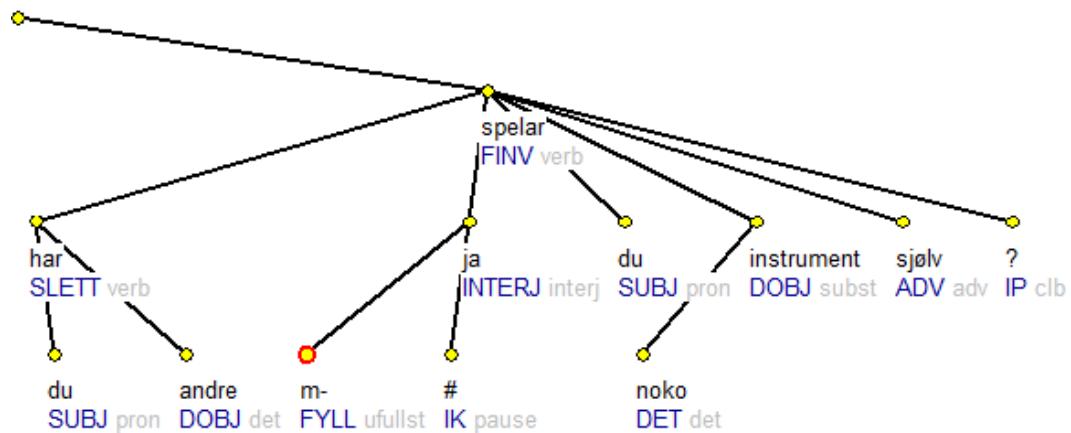
3.2 Restarts

SLETT
Eksempel: og alt ## når vi skulle ut så var det båten
 D & N s. 1568: parataxis: restart

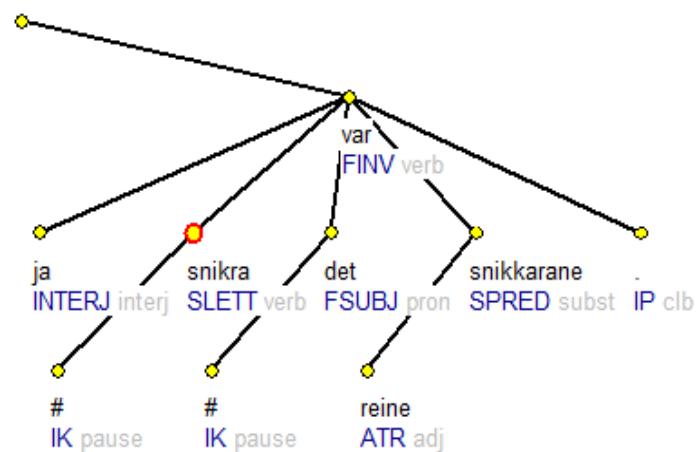
I LIA gir vi parataxis: restart-leddet funksjonen SLETT og setter det på høyeste node i det følgende leddet, eller på rotnoden, dersom det ikke finnes noe etterfølgende ledd.

Restarts skiller seg fra repairs når det foregående leddet ikke er relevant eller «beslektet» med det siste. Hvis en ytring starter med en ordsalat for så å gå over i en forståelig ytring, har vi å gjøre med en restart framfor en repair. En SLETT vil typisk ikke være en fullstendig syntaktisk enhet og mangler typisk et finitt verb.

(11) Har du andre m- # ja spelar du noko instrument sjølv ? (austevoll_uib_01 ... 0004 25/50)



(12) ja # snikra # det var reine snikkarane . (aamot_uio_01 ... 0000 149/423)

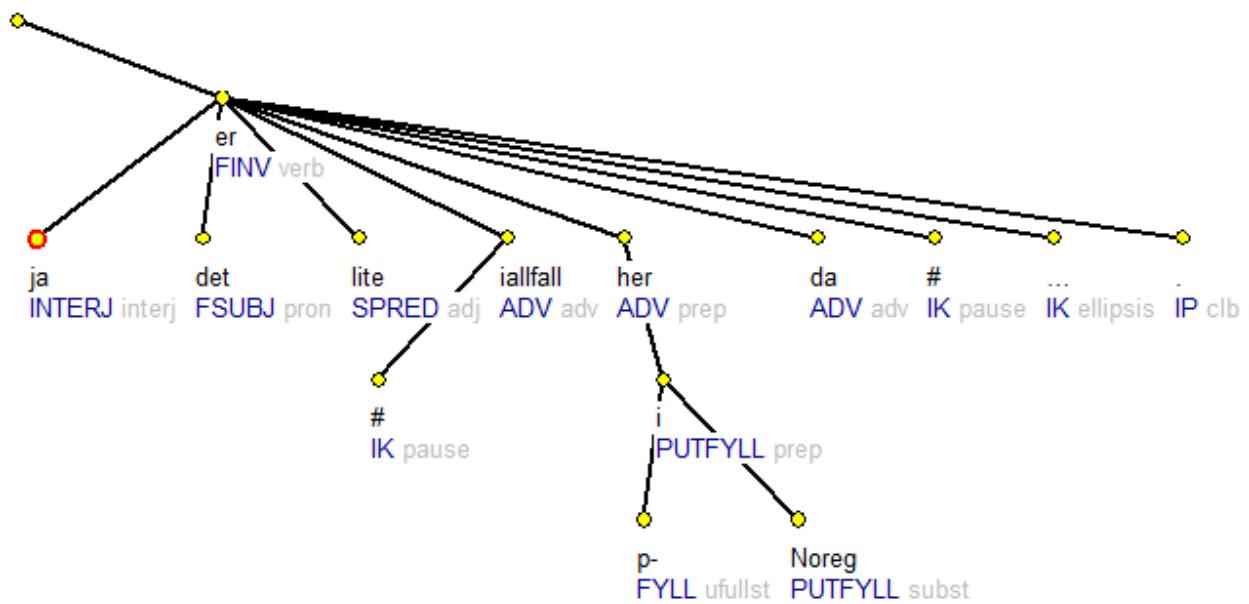


Se også kapittel 4, «Ufullstendige ytringer», for flere bruksområder for SLETT.

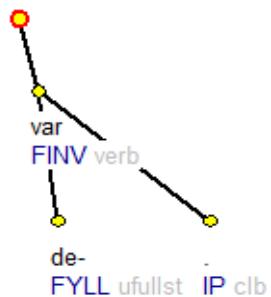
3.3 Ufullstendige ordstrenger

I transkripsjonene kan det forekomme strenger som avsluttes med en bindestrek (-). Dette er innført for at transkripsjonen skal kunne vise når talere avbryter et ord. Hvis dette avbrutte ordet ikke har noe åpenbart slektskap til noen av de nærmeste ordene, gis det funksjonen FYLL og behandles som pauser (se 2.2). Hvis det derimot formmessig ligner et nærliggende ord, kan det behandles som en repair (se 3.1).

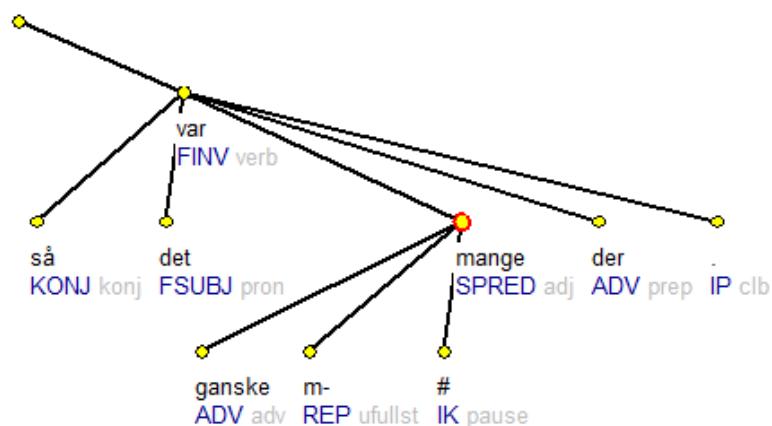
(13) ja det er lite # iallfall her p- i Noreg da # (austevoll_uib_01 ... 0001 48/50)



(14) var de- . (aamot_uio_01 ... 0000 68/423)



(15) så det var ganske m- # mange der . (aamot_uio_01 ... 0000 73/423)



4 Ufullstendige ytringer

Vi følger prinsippet i NDT om at det ikke finnes tomme noder. Eksempel: er subjektet utelatt, blir det ingen SUBJ-node.

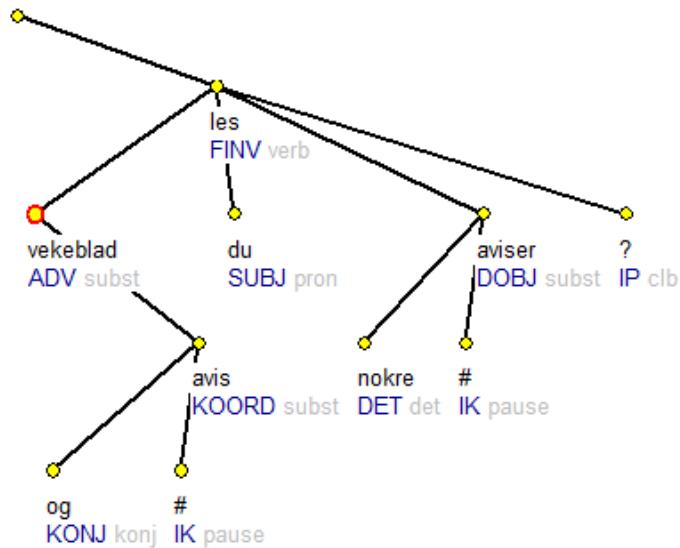
4.1 Ufullstendige ytringer i begynnelsen av et segment

Noen segmenter begynner med et setningsfragment som ikke inneholder et finitt verb og hvor det deretter kommer en setning med et finitt verb.

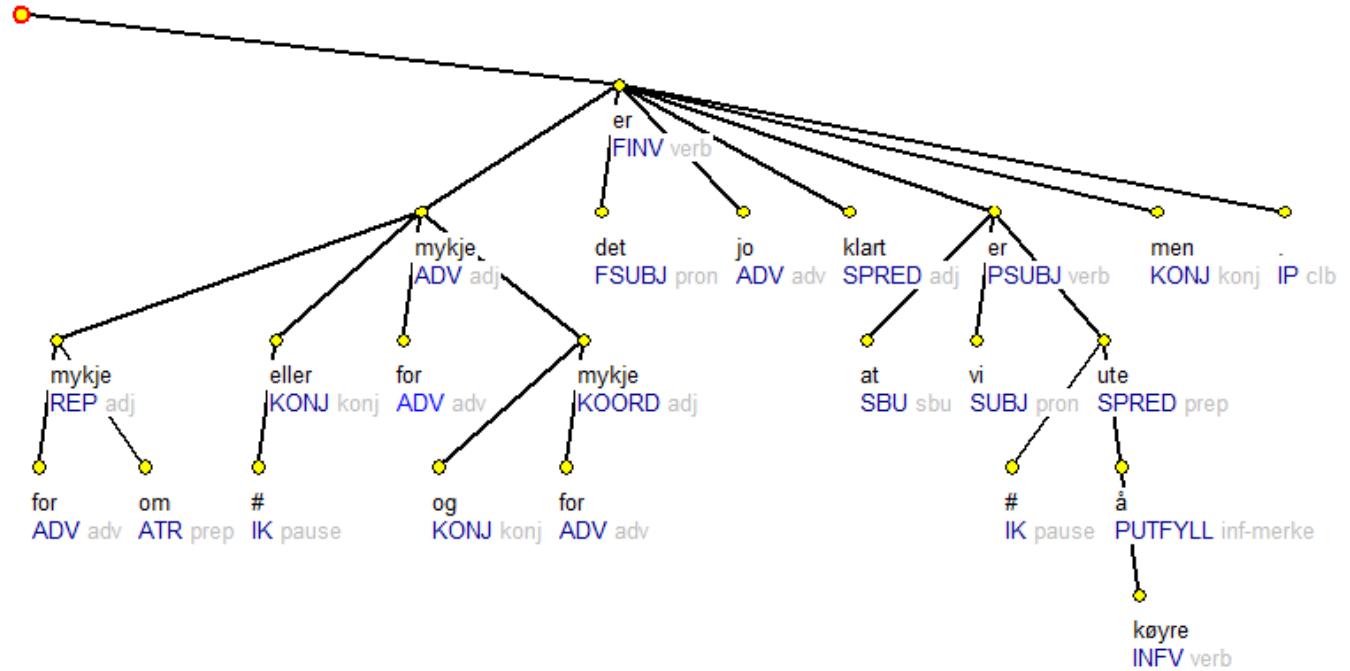
(16) vekeblad og # avis les du nokre # aviser? (austevoll_uib_01 ... 0003 12/50)

(17) for mykje om # eller for mykje og for mykje det er jo klart at vi er # ute å køyre men .
(austevoll_uib_01 ... 0005 34/50)

Denne første delen av setningen vil som oftest enten ha funksjonen ADV eller SLETT. En slags tommelfingerregel er at funksjonen ADV brukes hvis den ufullstendige ytringen passer semantisk sammen med det som kommer i resten av ytringen slik som i eksempel (16):



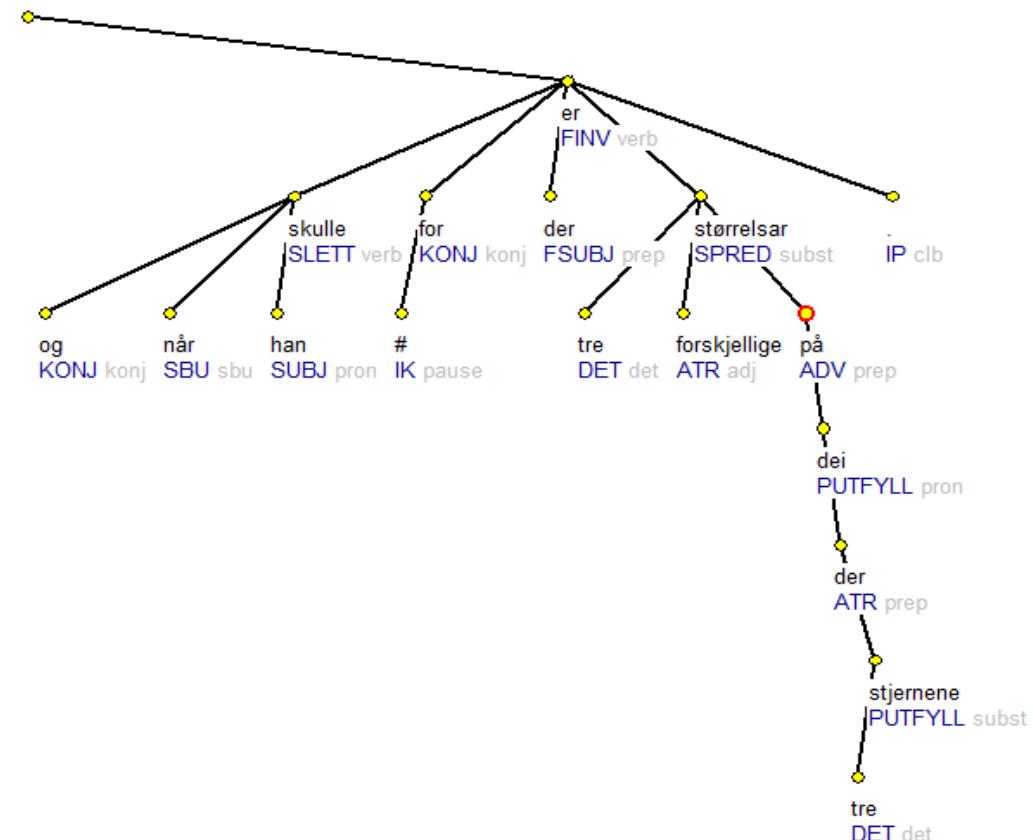
Og (17):



Hvis det ikke er noen semantisk sammenheng mellom den ufullstendige ytringen og det som kommer senere, brukes SLETT.

(18) Og når han skulle # for der er tre forskjellige størrelsar på dei der tre stjernene .

(flakstad_uib_04 ... 0021 35/50)

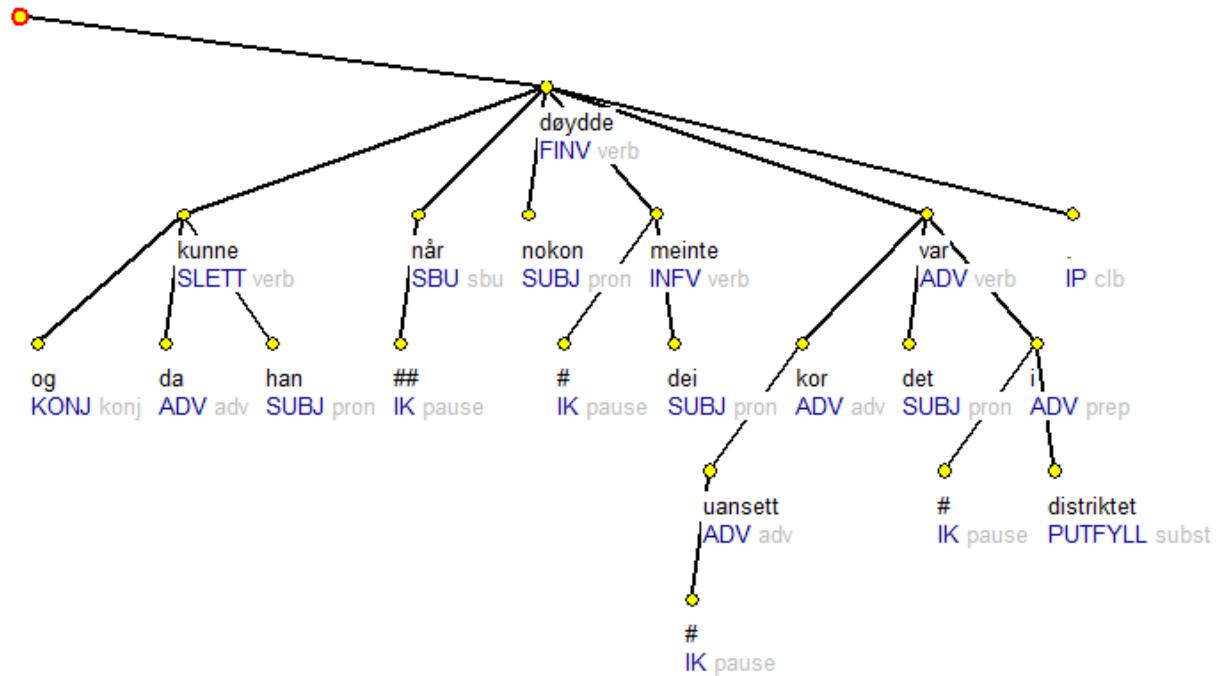


Hvis det er tydelig at det første som står i setningen er en repair skal det ha funksjonen REP som i eksempel (17). Les mer om REP i 3.1.

Fordi inndelingen av segmentene dessverre ikke alltid tar hensyn til setningsgrenser, vil noe som hører til forrige segment, noen ganger komme først i det neste segmentet. Det vil som oftest være naturlig å annotere dette som SLETT fordi det ikke har noen innholdsmessig sammenheng med det som kommer etter.

(19) og da kunne han ## når nokon døydde # meinte dei # uansett kor det var # i distriktet .

(flakstad_uib_04 ... 0000 19/50)

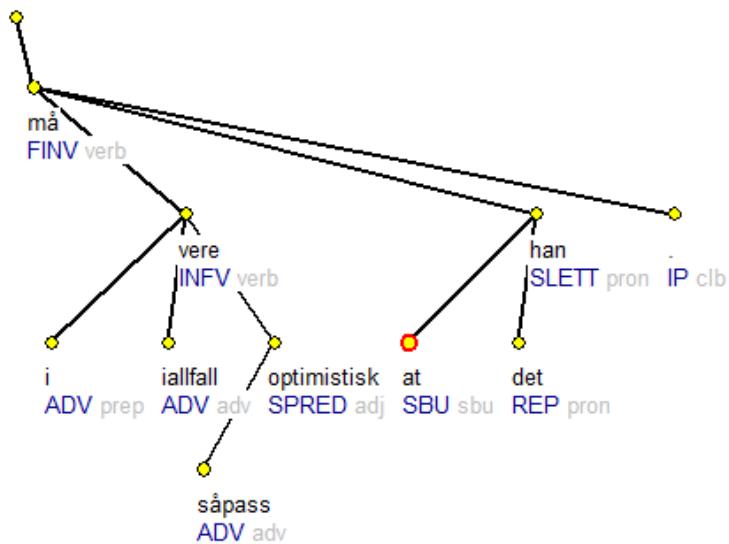


Disse segmentene har vist seg å være vanskelig å annotere og derfor gi en enkel løsning. Ofte må vi bare bruke skjønn når vi møter på slike segmenter.

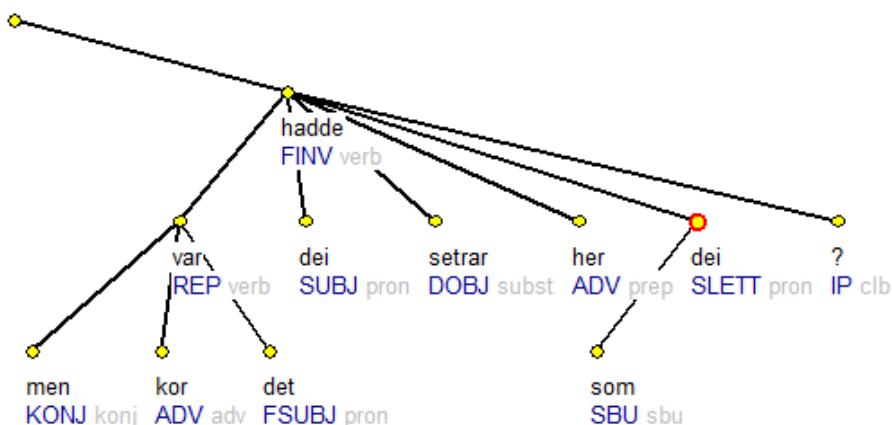
4.2 Ufullstendige ytringer i slutten av et segment

Det forekommer at et segment avsluttes med en ufullstendig ytring; denne kan bestå av enkeltord eller større strukturer. Denne typen ytringer er dependent på rotnoden og har funksjonen SLETT.

(20) må i iallfall vere såpass optimistisk **at det han** . (giske_uib_02 ... 0015 9/50)



(21) men kor var det hadde dei setrar her **som dei** ? (bolsoey_uio_03 ... 294/250)



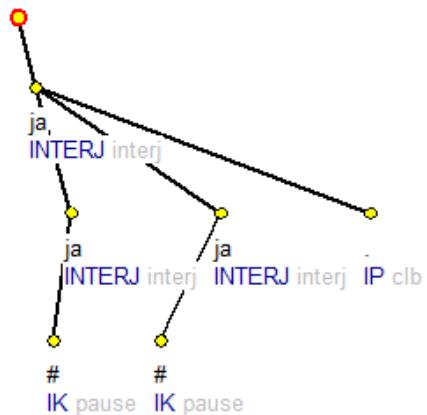
4.3 Ytringer uten finitt verbal

Ytringer uten finitt verbal behandles som fragmenter slik som i NDT. Det er to typer, den ene får funksjonen INTERJ, den andre FRAG. En interjeksjon er alltid å foretrekke som rot. Det vil si, hvis en har en interjeksjon i segmentet, settes den til rotnode. FRAG nytter vi først og fremst som en siste mulighet.

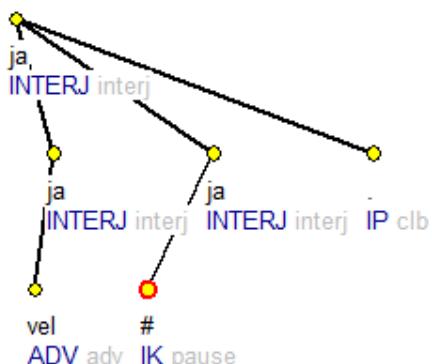
4.3.1 INTERJ:

Det er ganske vanlig at et segment bare inneholder interjeksjoner. Da skal den første interjeksjonen være hode, og de andre henge på denne interjeksjonen. Alle interjeksjoner skal ha funksjonen INTERJ, men husk at segmentet kan inneholde andre typer ledd som er dependent på rota; det kan være andre interjeksjoner, adverb, substantiv, infinitiv verb m.m..

(22) ja # ja # ja . (aamot_uio_01 ... 0000 139/423)

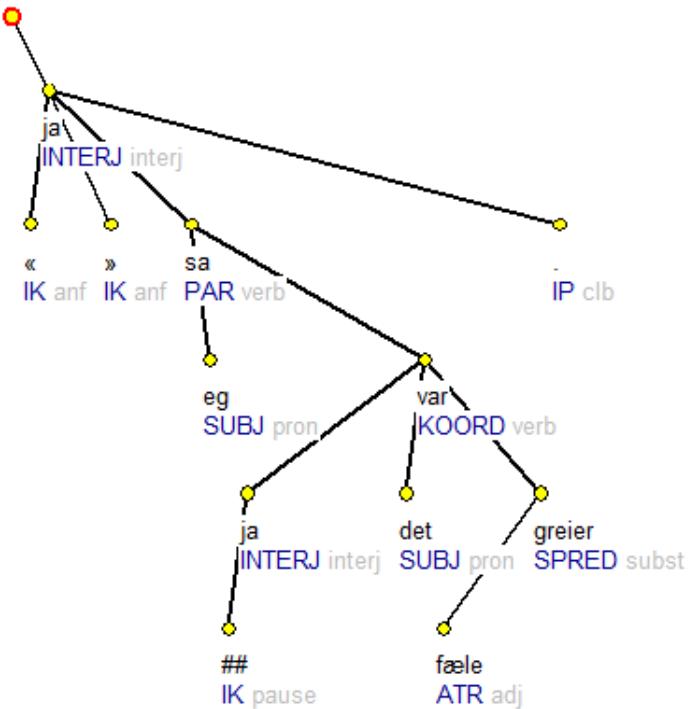


(23) ja vel ja # ja . (trondheim_uib_10 ... 0000 17/216)



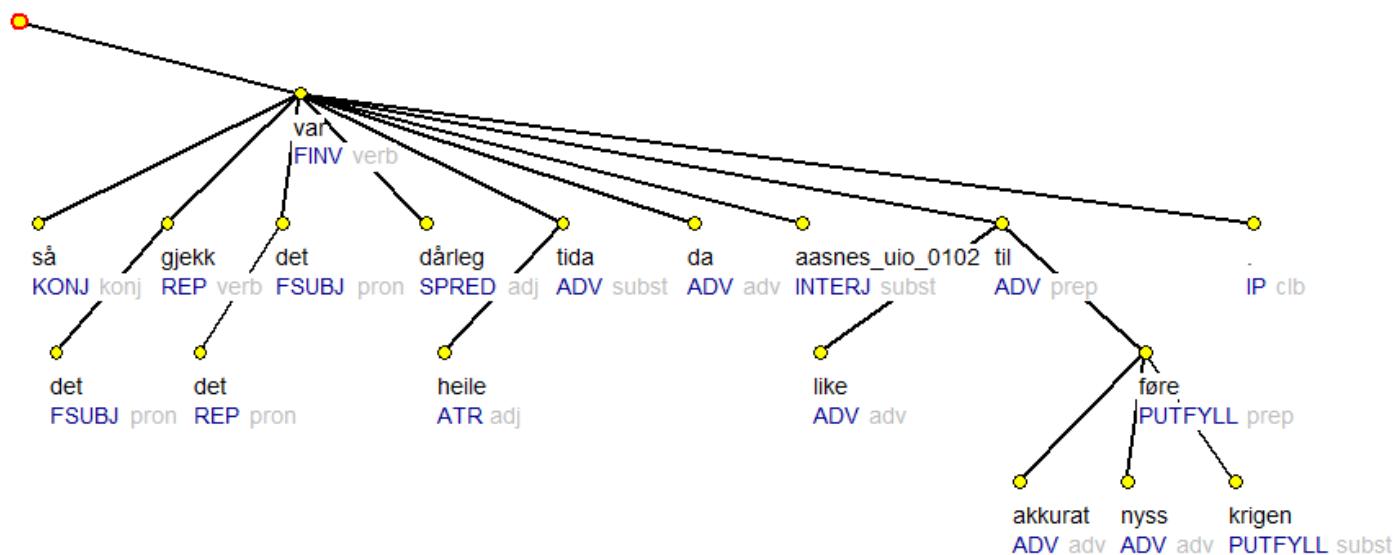
På grunn av måten vi behandler direkte tale (se 6.1, og NDT 5.3), kommer man iblant over tilfeller som (24), hvor INTERJ skal være rota i segmentet, til tross for at det inneholder finitt verb.

(24) «ja» sa eg ## ja det var føle greier . (aamot_uio_01 ... 0000 16/423)



I tillegg kan også egennavn (i mange tilfeller navnekoder, i dette prosjektet) og pronomener fungere som interjeksjon:

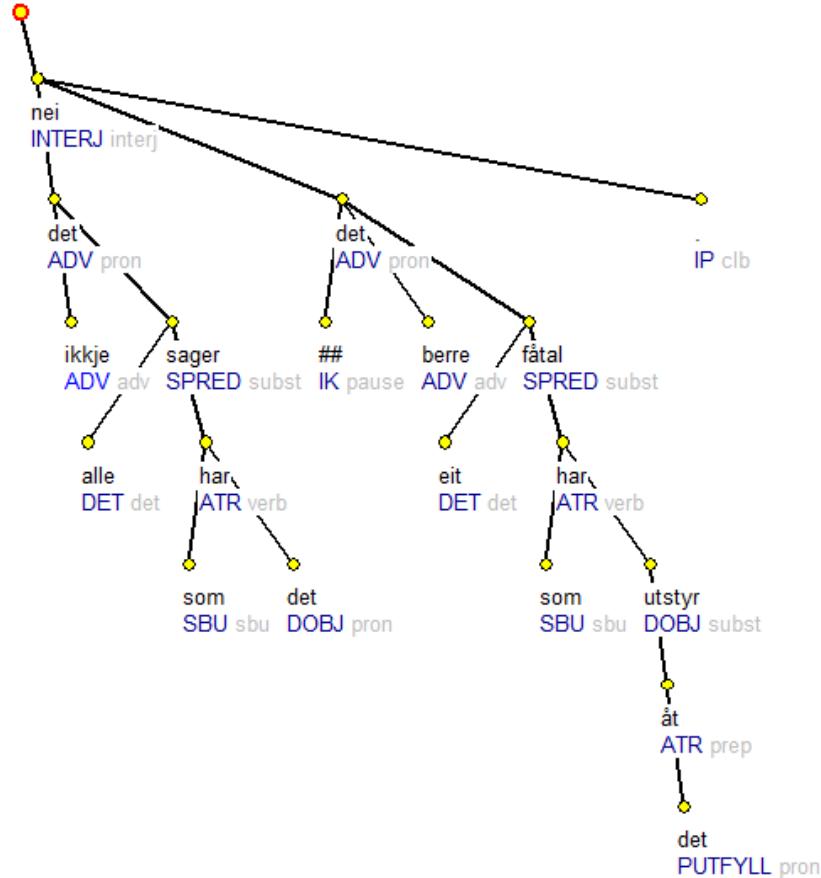
- (25) så det gjekk det det var dårlig heile tida da **aasnes_0102** like til akkurat nyss føre krigen .
 (aasnes_01 ... 460/466)



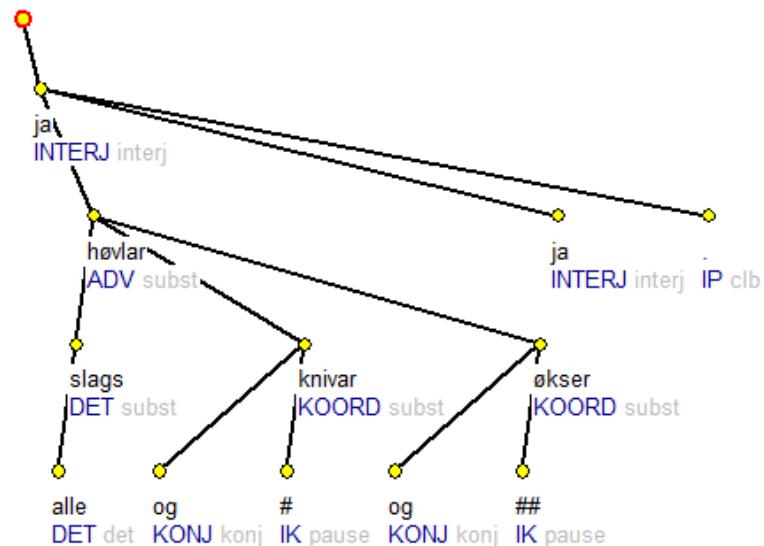
En vil ofte komme over segmenter med INTERJ-rot som også inneholder større dependensstrukturer som kan ha ulike typer ledd (unntatt finitt verb) som hode. Ofte vil dette ledet være et nominal.

Substantiver, adjektiver, adverb, preposisjoner og infinitte verb som er dependent på en INTERJ-rot, men hødet til en større struktur i samme segment, får funksjonen ADV.

- (26) nei det ikkje alle sager som har det ## det berre eit fåtal som har utstyr åt det .
 (bolsoey_udio_03 ... 37/250)²



- (27) ja alle slags høvlar og # knivar og ## økser ja . (bolsoey_udio_03 ... 91/250)



² For mer om ellipser, se kapittel 10.

4.3.2 FRAG:

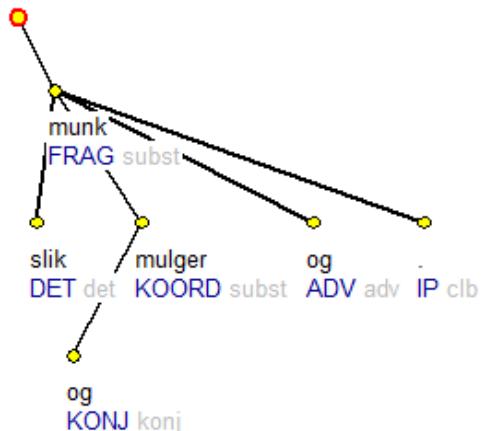
En støter ofte på segmenter uten verken finitt verbal eller interjeksjon. I slike tilfeller har vi ikke noe annet valg enn å sette funksjonen FRAG på rot-noden. Det kan være vanskelig å avgjøre hva som skal være det øverste hodet i slike ytringer. Vi følger *prominenshierarkiet* gitt i NDT (s. 119) for å avgjøre dette.

- (28) infinitt verbal > subjekt > indirekte objekt, subjektspredikativ > potensielt subjekt, direkte objekt > objektspredikativ > adverbial

Dersom prominenshierarkiet ikke er tilstrekkelig for å avgjøre hva som er hode, se NDT (s.27) for mer om FRAG.

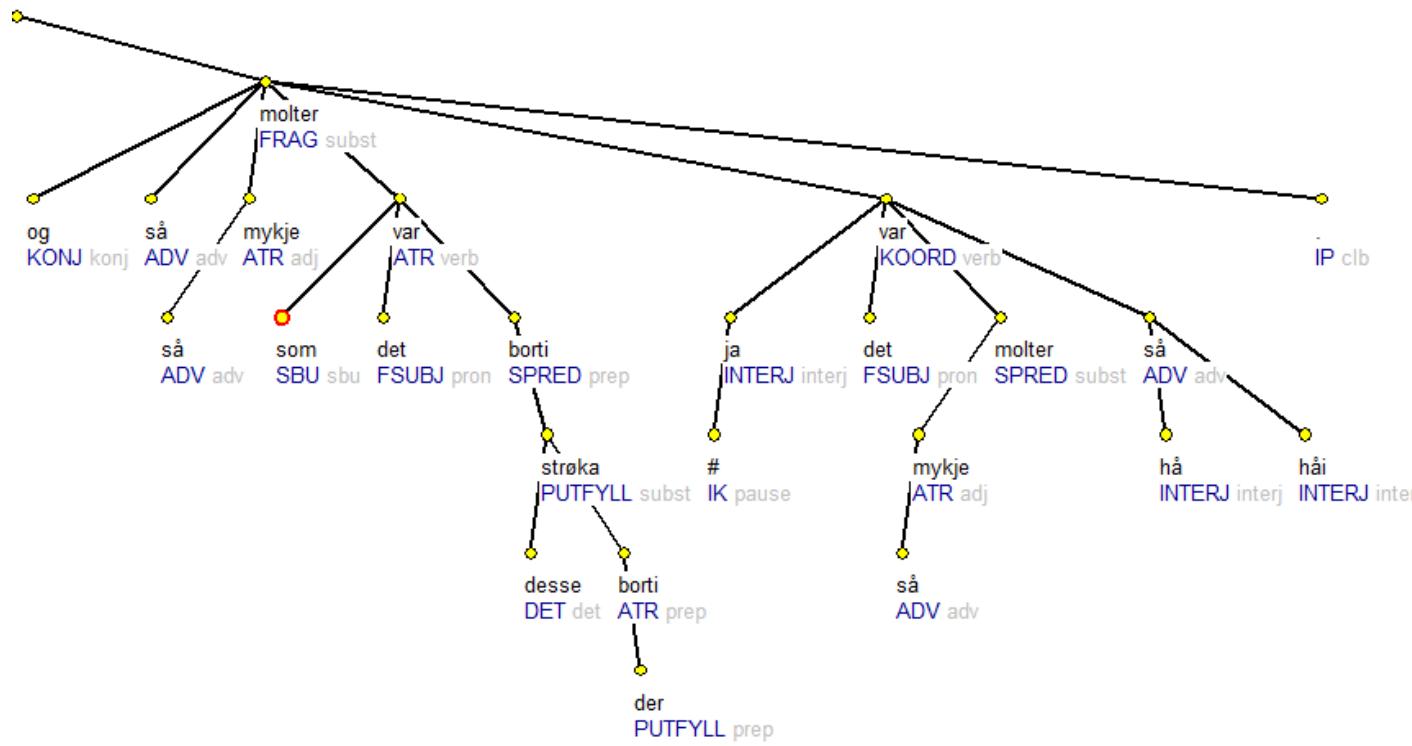
Et eksempel på et segment med en klar FRAG-løsning.

- (29) slik munk og mulger og . (aamot_uio_01 ... 0000 76/423)



I segmenter som (30), må vi også sette rot-noden som FRAG, til tross for at det inneholder finitt verbal. Dette er fordi «molter» her fungerer som korrelatet til en følgende relativsetning, og korrelatet må være hodet til det relativiserte ledet.

(30) og så så mykje molter som det var borti desse strøka borti der # ja det var så mykje molter så
hå håi . (aamot_uio_01 ... 0000 26/423)



Det er vanskelig å gi helt klare regler på når FRAG skal benyttes, så man er ofte nødt til å bruke skjønn (se NDT kap 5.18 for flere råd).

5 Discourse elements

Interjeksjoner og andre responsord behandler vi også som i NDT.

5.1 Interjeksjoner

Interjeksjoner får funksjonen INTERJ og har enten rotfunksjon eller er typisk dependent på et verb.
I segmenter som bare inneholder en interjeksjon, har interjeksjonen funksjonen INTERJ (se også 4.3.1.).

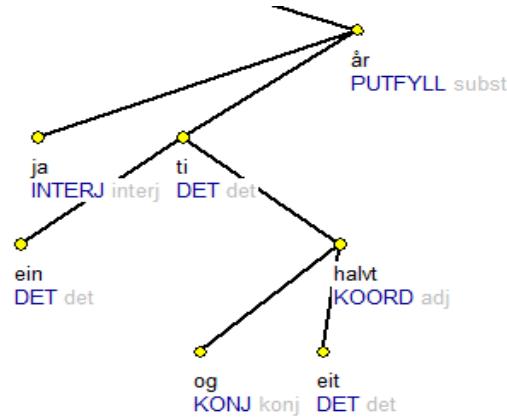


Det er også mulig å ha en interjeksjon i en større ytring slik som i eksempel (31).

Denne interjeksjonen skal henge på den høyeste noden i ytringen der den hører hjemme.

(31) e ja s- eg har budd i lla i e åtte år # og før den tid så budde eg oppi Klæbu i **ein ja ti og eit**

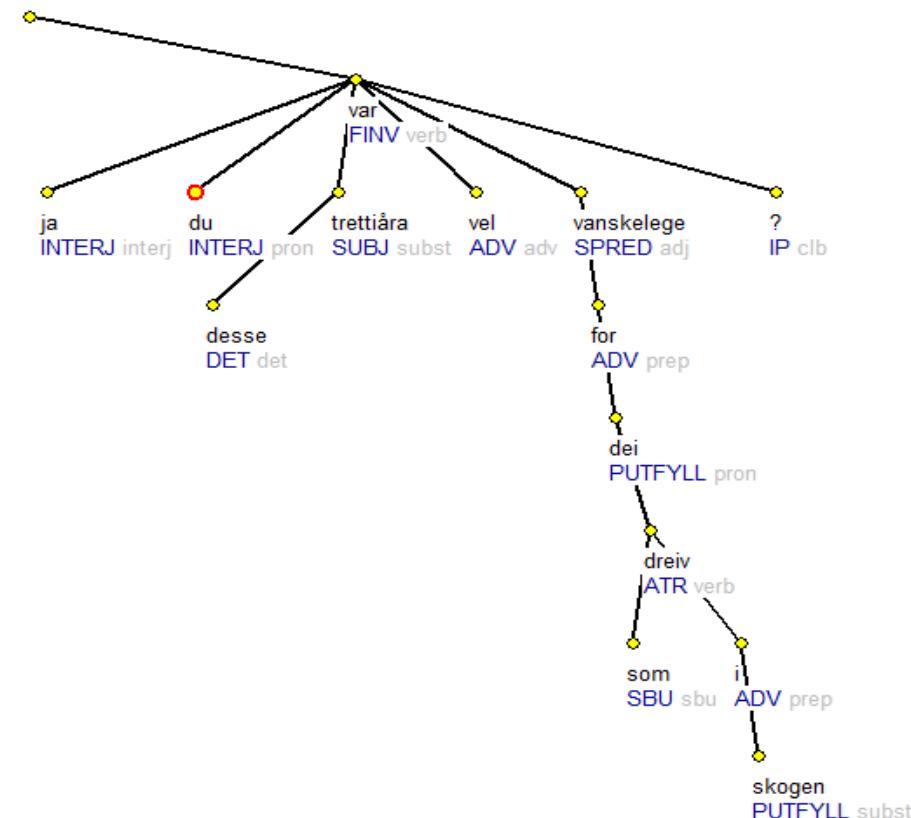
halvt år cirka . (trondheim_uib_10 ... 0000 6/216)



Husk at flere ord som ikke vanligvis blir klassifisert som interjeksjoner, kan brukes som interjeksjoner i talemål slik som *du* i eksempel (32). Det er ikke mulig å lage en uttømmende liste dessverre.

(32) ja du desse trettiåra var vel vanskelege for dei som dreiv i skogen ? (aamot_udio_01 ... 0000

306/423)



5.2 Discourse fillers

Nølelyder som *e* og *m* (discourse filler i Dobrovolje & Nivre) gir vi ordklassetaggen *nol* og funksjonen FYLL. Nølelydene behandler vi likt som pausetegnet og er, dermed, dependent til ordet som kommer til høyre i setningsstrengen.

FYLL

(33) nei å jo dei går på inn- innpå e fastlandet .

FYLL



(34)e ja her oppi e der er vel mykje hare ?

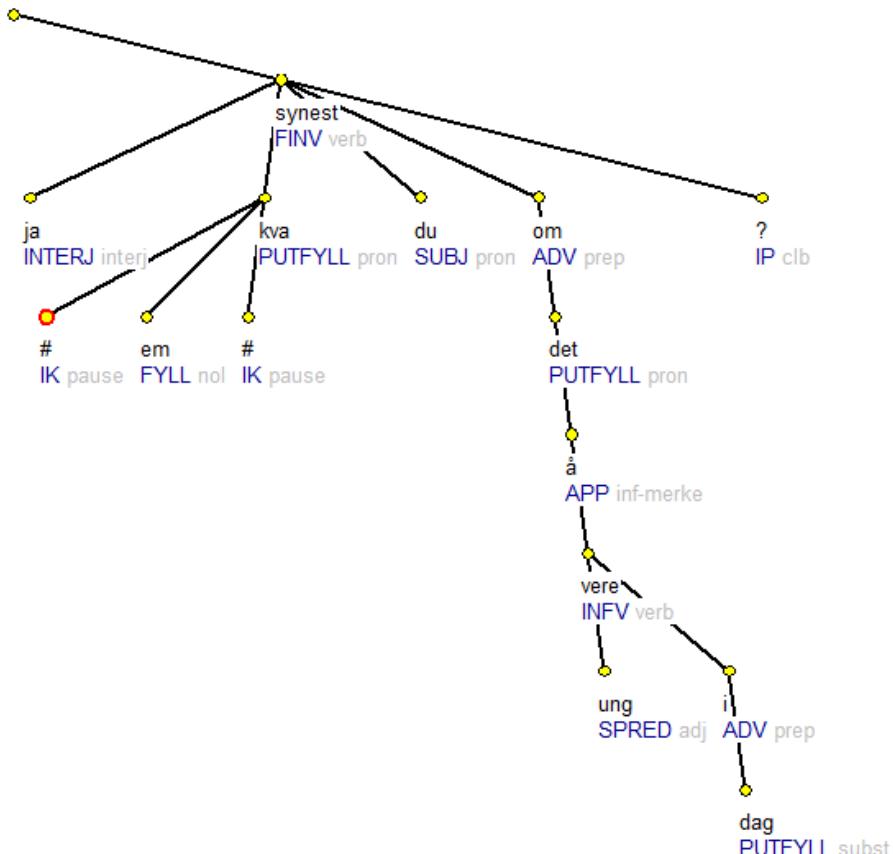
FYLL



(35) ja der kom så e tidleg gras om våren da

Hvis det kommer både nølelyder og pauser før et ord i strengen henger alle disse på ordet til høyre i setningsstrengen. Det skal ikke henge pauser på nølelyder eller omvendt.

(36) ja # em # kva synest du om det å vere ung i dag? (austevoll_uib_01 ... 0005 18/50)

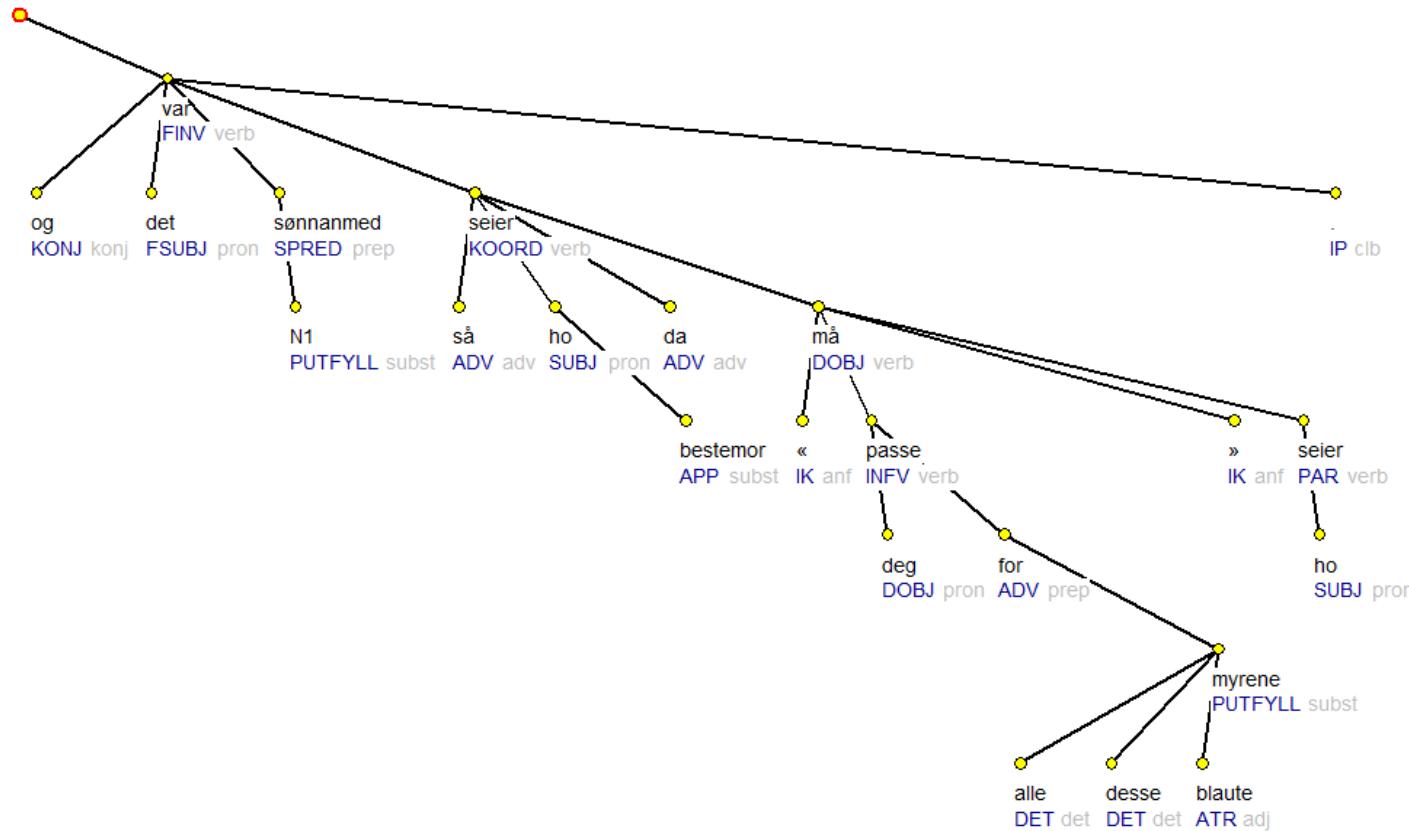


6 Parentetiske uttrykk

6.1 Direkte tale

Utsagnsverb ved direkte tale skal ha funksjonen PAR, dependent på den relevante helsetningens hode. Se NDT 5.3.3.

(37) og det var sønnanmed N1 så seier ho da bestemor «må passe deg for alle desse blaute myrene» seier ho. (aamot_udio_01 ... 0000 8/423)



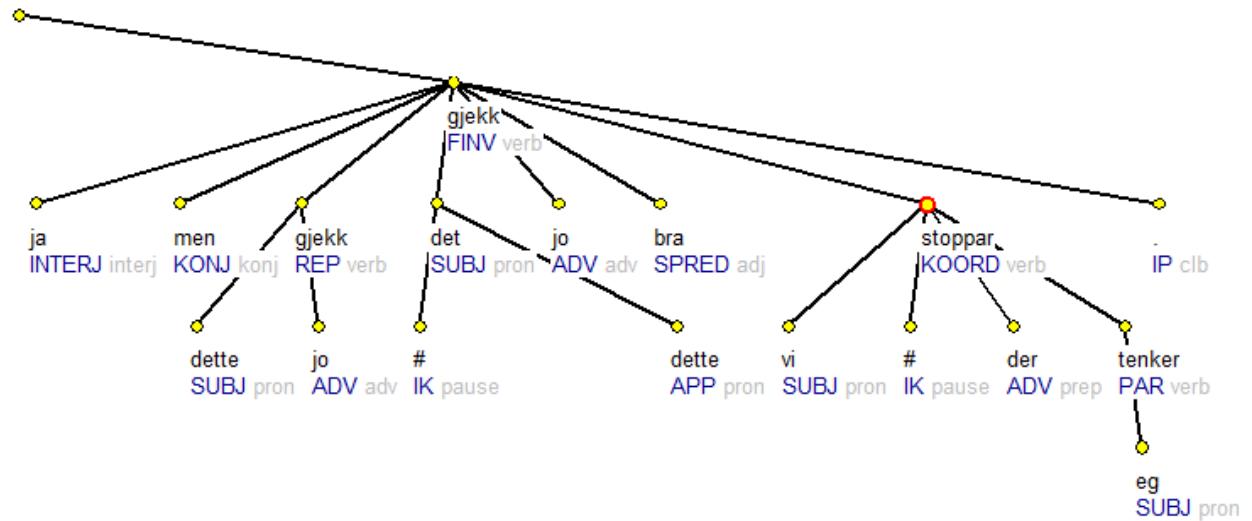
For hvordan man annoterer anførselstegn ved direkte tale, se 2.2.2.

6.2 Andre typer utsagnsverb

Vi annoterer andre typer utsagnsverb med PAR også. I denne gruppen blir «veit du», «trur eg», «tenker eg», «syntes eg» og tilsvarende utsagn inkludert. Disse fungerer ikke på samme måte som andre finitte verb, ved at de har en fast ordstilling (alltid subjektet til slutt) og de knytter ikke til seg andre obligatoriske ledd enn subjektet. Dessuten kan disse utsagnsverbene komme på andre plasser i en ytring sammenlignet med andre finitte verb. Denne typen utsagnsverb er, likt som utsagnsverb ved direkte tale, dependent på hodet til den relevante helsetningen.

(38) ja men dette gjekk jo # det gjekk jo bra dette vi # stoppar der **tenker** eg .

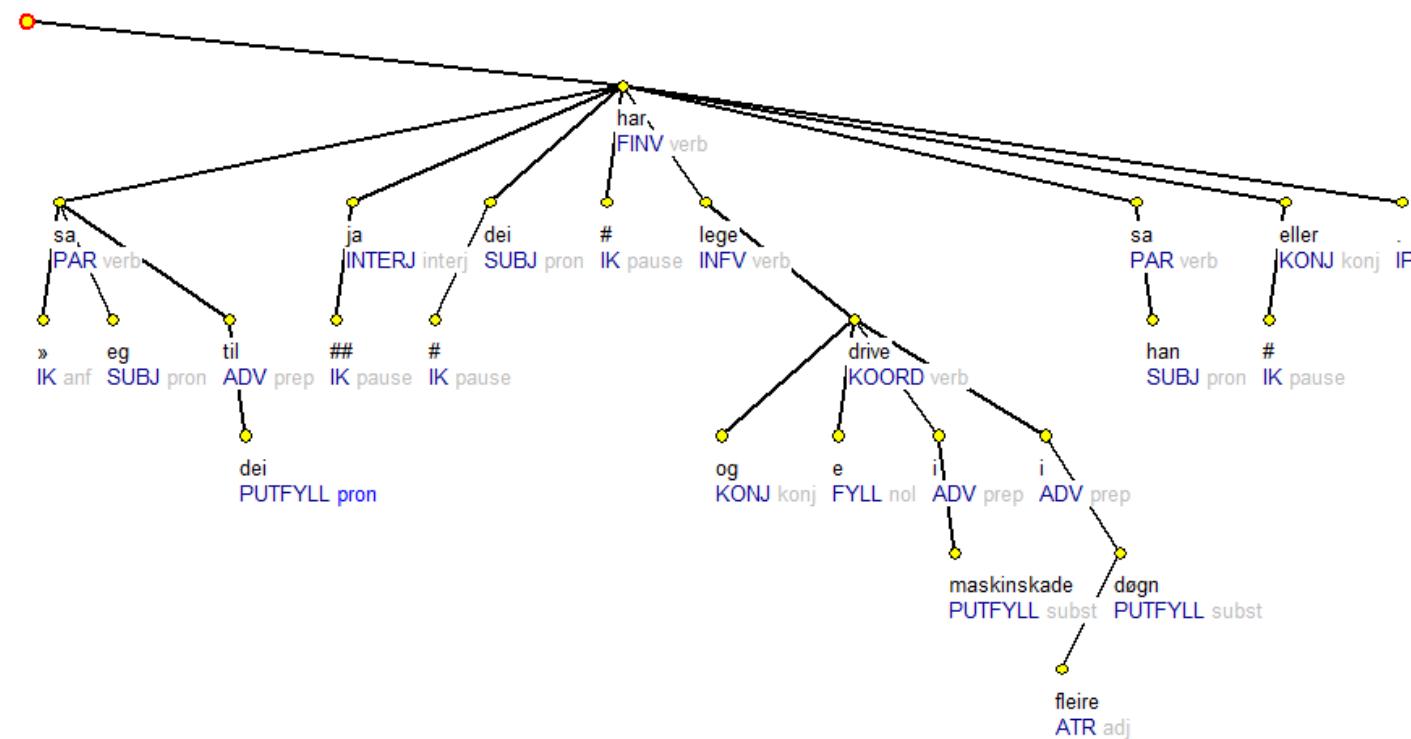
(austevoll_uib_01 ... 0007 1/20)



Vi finner også parentetiske uttrykk i starten av segmenter. Disse leddene tilhører typisk det foregående segmentet, og er ikke tett knyttet til det inneværende segmentets øvrige innhold på noen åpenbar måte, og får derfor funksjonen PAR og er dependent på *rotnoden*; dette fordi at ledet ikke har et direkte relevant helsetningshode å knytte seg til.

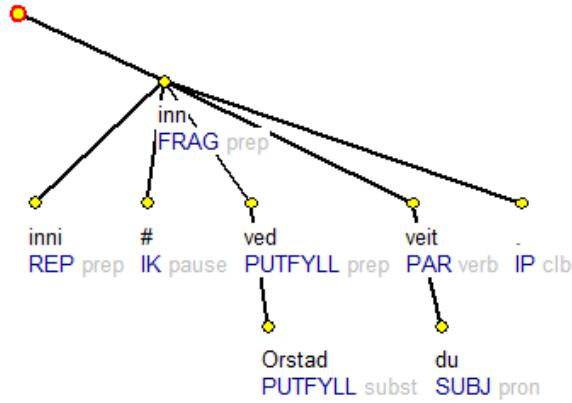
(39) « sa eg til dei ## ja # dei # har lege og e drive i maskinskade i fleire døgn sa han # eller .

(vardoe_udio_01 ... 0008 50/50)



Denne typen utsagnsverb kan også være dependent på INTERJ- og FRAG-røtter i visse tilfeller. Dette regnes som unntak til hovedregelen om at PAR skal være dependent på hodet til den relevante hovedsetningen.

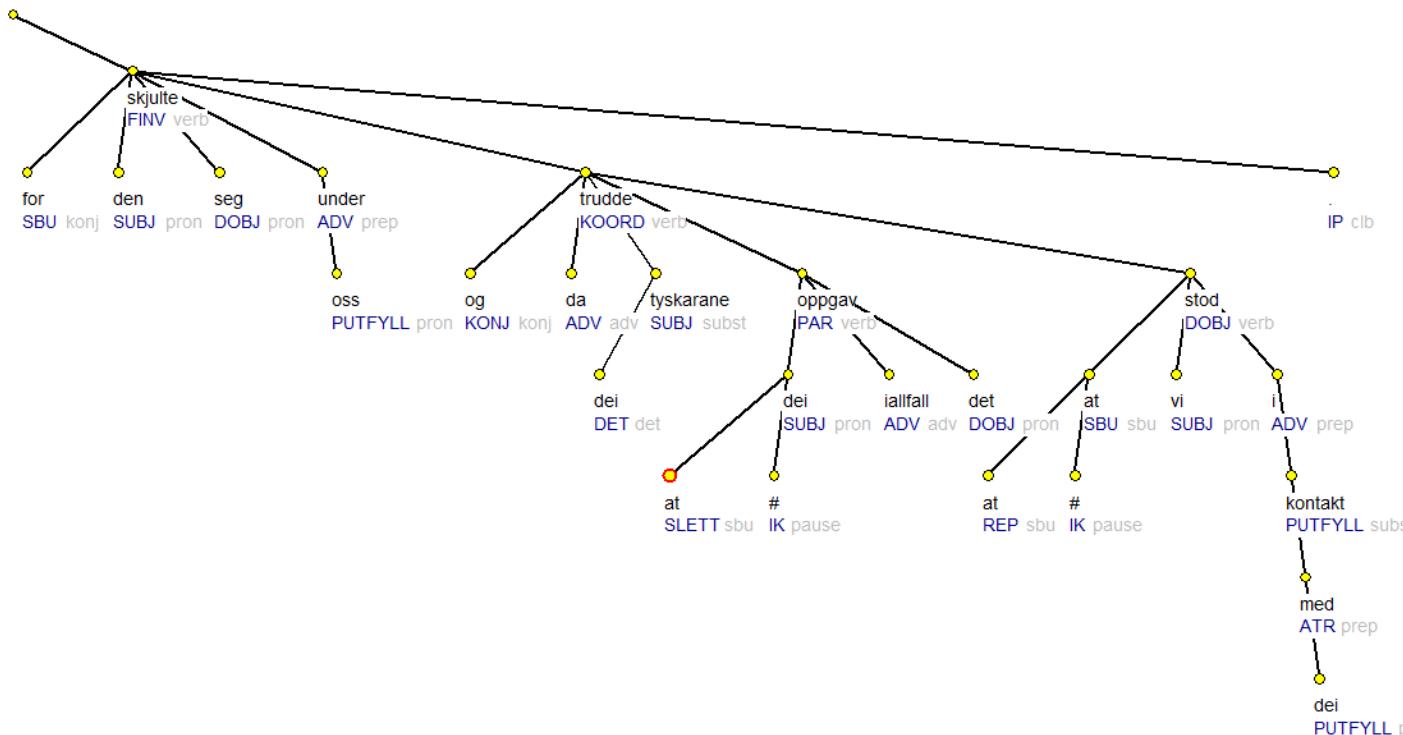
(40) inni # inn ved Orstad veit du . (eidsberg_udio_03 ... 0007 34/50)



6.3 Metakommentarer

Kommentarer fra samtaledeltakerene om eget utsagn eller diskursen, skal være PAR og dependent på hodet til den relevante hovedsetningen:

(41) For den skjulte seg under oss og da trudde dei tyskarane at # **dei oppgav ialfall det** at # at vi stod i kontakt med dei . (vardoe_udio_01 ... 0007 21/50)



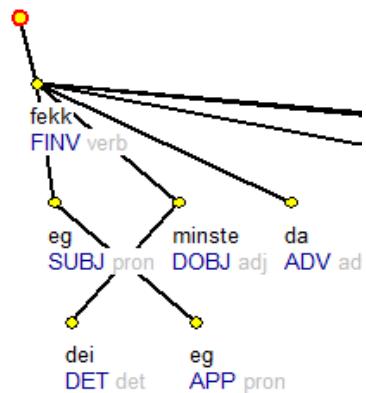
7 Apposisjoner

Se NDT 4.2 APP

7.1 Subjektsdublering

Det er ganske vanlig med subjektsdublering. Da har man gjerne et subjekt på en av de vanlige subjektsplassene (før og etter det finitte verbet) og i tillegg en dublering av subjektet på slutten av setningen. Subjektet som står på den vanlige subjektsplassen får funksjonen SUBJ, mens subjektsdubleringen på slutten av setningen får funksjonen APP og er dependent på SUBJ.

(42) fekk **eg** dei minste **eg** da (aamot_uio_01 ... 0000 119/423)

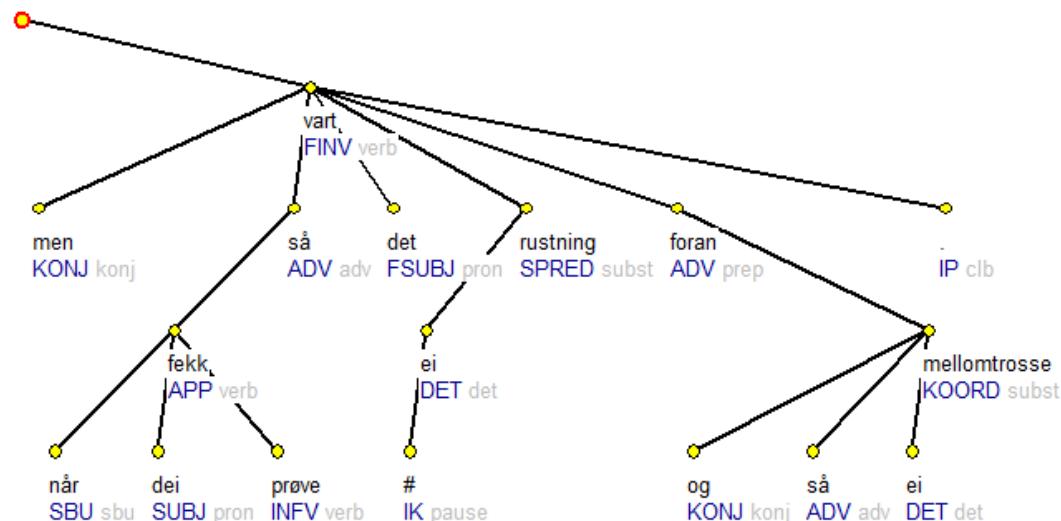


7.2 APP når et foranstilt ledds innhold gjentas med *så* eller *da* (4.2.4. i NDT)

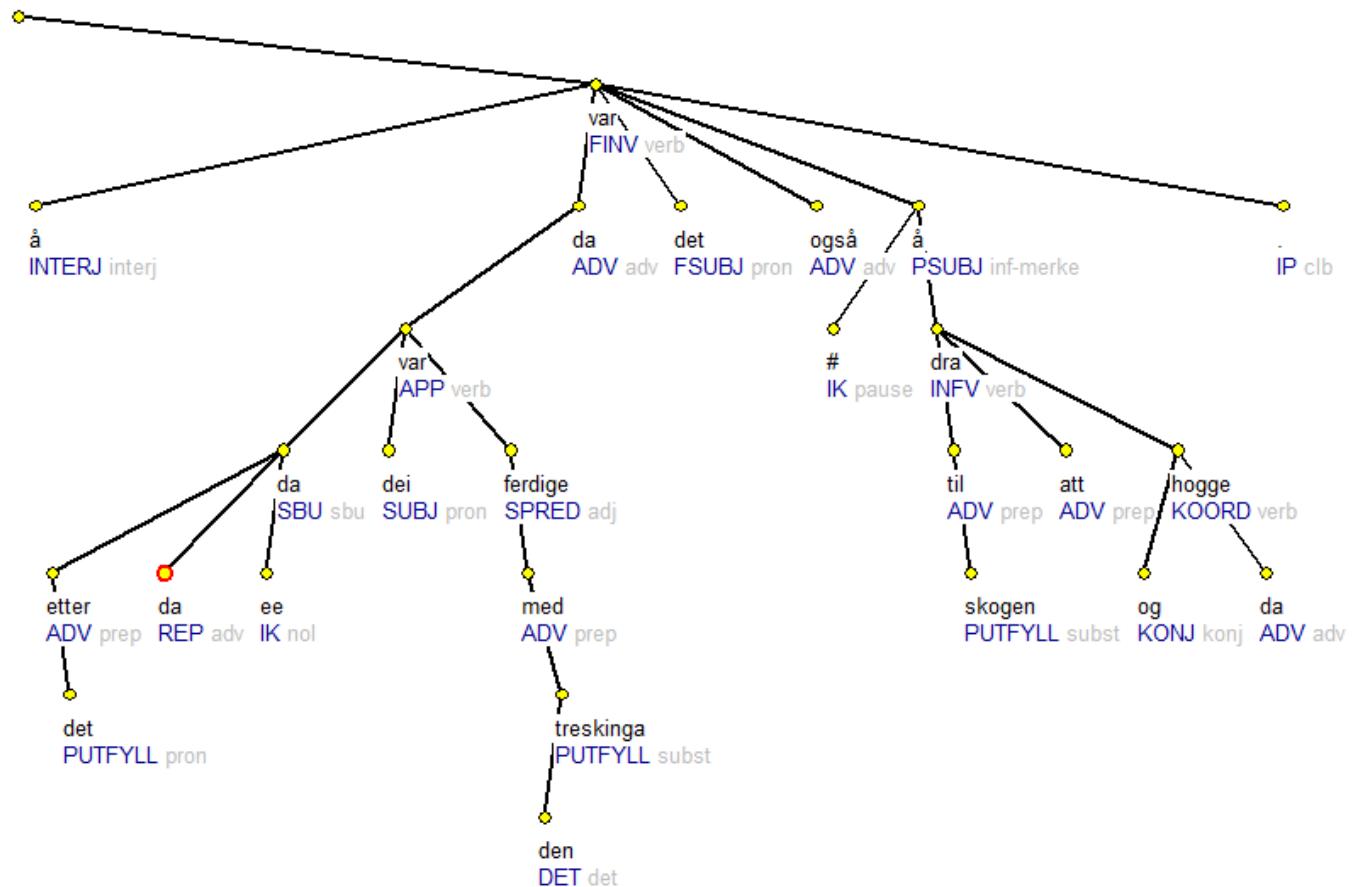
Det er vanlig at et foranstilt, adverbielt ledd følges av et *så* eller *da*, som videre følges av det finitte verbet:

(43) men **når** **dei** **fekk** **prøve** **så** **vart** **det** # **ei** **rustning** **foran** **og** **så** **ei** **mellomtrosse**

(aamot_uio_01 ... 0000 203/423)



(44) å etter det da ee **da** **dei** var ferdige med den treskinga da var det også # å dra til skogen att
og hogge da . (aamot_uio_01 ... 0000 224/423)



8 Formelle subjekter

Det er vanskelig å skille mellom referensielle subjekter og formelle subjekter. Selv om vi i stor grad bare forholder oss til innholdet i det segmentet vi annoterer, må vi bruke konteksten for å avgjøre om noe er et formelt subjekt eller et referensielt subjekt.

Husk at et formelt subjekt ikke kan ha en referensiell APP. Men et formelt subjekt kan ha et innholdstømt *det* som APP.

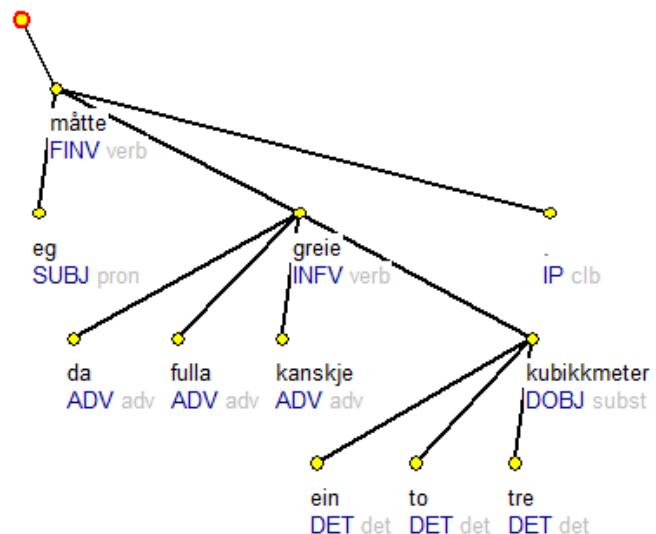
Les mer om formelle subjekter i NDTs retningslinjer (3.2.2).

9 FINV og INFV

I konstruksjoner med hjelpeverb, skal hjelpeverbet være hodet, mens det leksikalske verbet skal være dependent med funksjonen INFV.

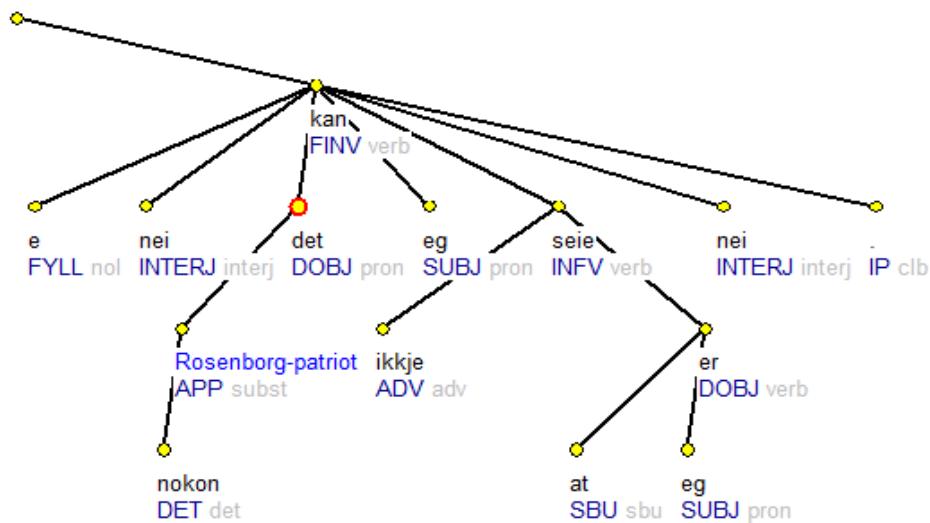
I følge NDT (3.1.4) skal konjunksjoner, interjeksjoner, subjunksjoner, setningsskilletegn, setningsinterne skilletegn, subjekter og formelle subjekter være dependenter på det finitte hjelpeverbet, mens potensielle subjekter og alt annet skal være dependenter på det leksikalske verbet. Det betyr at for eksempel alle ADV skal henge på INFV og ikke på det finitte hjelpeverbet.

(45) eg måtte da fulla kanskje greie ein to tre kubikkmeter (aamot_uio_01 ... 0000 50/423)



(46) e nei nokon Rosenborg-patriot det kan ikkje eg seie at eg er nei . (trondheim_uib_10 ... 0000

50/216)



10 Ellipser

Det er ikke uvanlig med ellipser i talespråk, særlig av verb og substantiv. Vi følger i stor grad NDT i behandlingen av ellipser (NDT kap 5.1), og forholder oss til prominenshierarkiet (se også 4.3.2 her). På grunn av måten vi har segmentert filene i LIA-prosjektet, ser vi iblant ellipser av helsetnings- og leddsetningsverb inne i segmenter. I slike tilfeller skal det mest prominente argumentet til det eliderte verbet fungere som hode, med funksjonen verbet «skulle hatt» i ytringen. Eventuelle tilhørende ledd til det eliderte verbet, er dependente på det mest prominente argumentet, som vanlig. For analysen av eliderte substantiv, se NDT kapittel 5.1.

11 Setningsknuter og preposisjonsstranding

I dette avsnittet omhandler vi setningsknuter og preposisjonsstranding sammen. Når et ledd fra undersetningen står som et ledd i oversetningen, kaller vi det for en setningsknode. Leddet som står i oversetningen kan enten være et subjekt (47), objekt (48) eller predikativ (49) fra leddsetningen. Ved preposisjonstranding står ledet som er utfylling til en preposisjon på en annen plass enn etter preposisjonen (gjerne i forfeltet før det finitte verbet.), slike som i eksempel (50) og (51).

(47) Mjølceforsyninga har ma- # trur eg har mange hestar i dag. (eidsberg_ ui0_03 ... 0003 34/50)

(48) aviser trur eg les kvar dag (austevoll_uib_01 ... 0003 14/50)

(49) Heilt ny vil eg ikkje seie han er. (austevoll_uib_01 ... 0003 15/50)

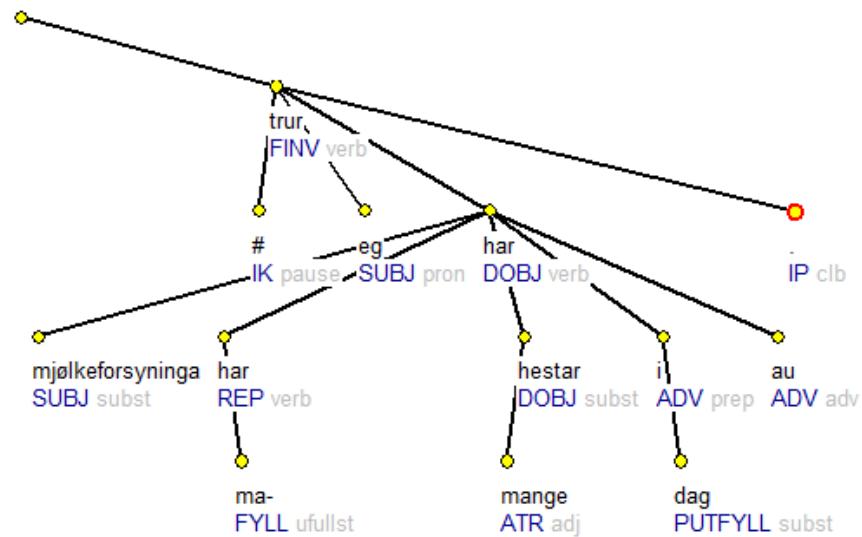
(50) Den var det ingen som hadde maken til. (eidsberg_ ui0_03 ... 0007 2/50)

(51) Generator måtte vi kjøre med under krigen. (eidsberg_ ui0_03 ... 0007 14/50)

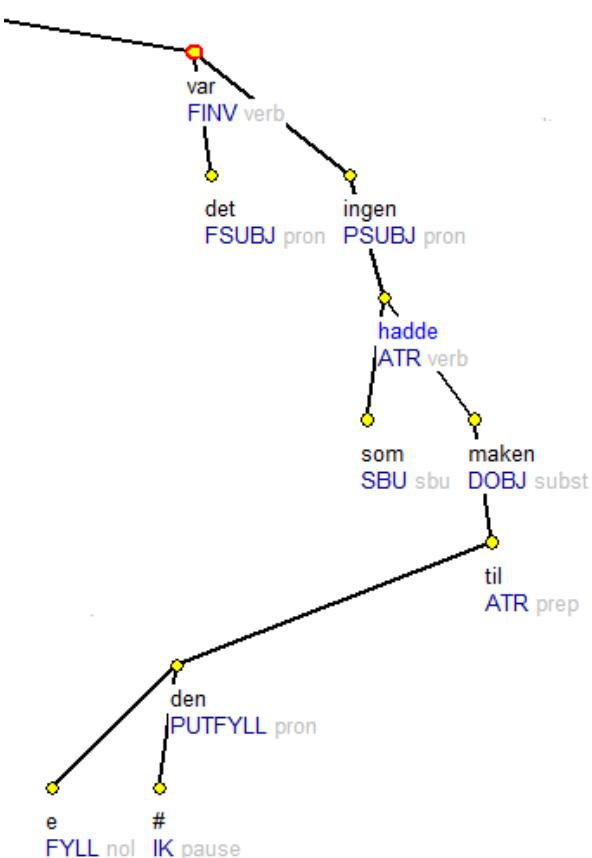
Setningsknuter dukker ofte opp i materialet og det er ikke så lett å finne ut hvor det står noe om setningsknuter i NDTs retningslinjer. I følge Per Erik Solberg (p.c.) skal det ledet som står i oversetningen eller som ikke lenger står etter preposisjonen, være dependent på verbet eller preposisjonen det semantisk tilhørte. Altså i eksempel (47) skal mjølceforsyninga være SUBJ på «har» og «heilt ny», i eksempel (49), skal være dependent på «er» og ha funksjonen SPRED. I eksempel (50) skal «den» være PUTFYLL og henge på «til».

Det står om denne typen konstruksjoner i NDTs retningslinjer, avsnitt 2.2.3. Det er viktig å påpeke at disse konstruksjonene er annerledes enn det som står beskrevet i NDTs retningslinjer, avsnitt 5.12.

Eksempel 47:



Eksempel 50:

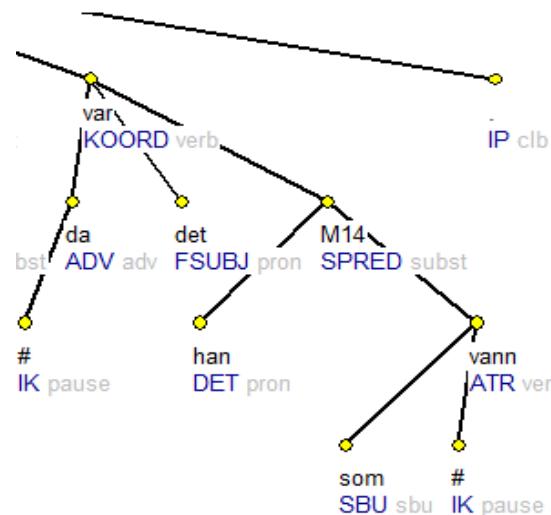


12 Andre frekvente talemålstrekk

12.1 Prepropriell artikkelen:

Prepropriell artikkelen skal ha funksjonen DET og være dependent på egennavnet eller substantivet det står til.

(52) da var det **han M14** som # vann (aamot_udio_01 ... 0000 253/423)



12.2 Konjunksjon på slutten av et segment

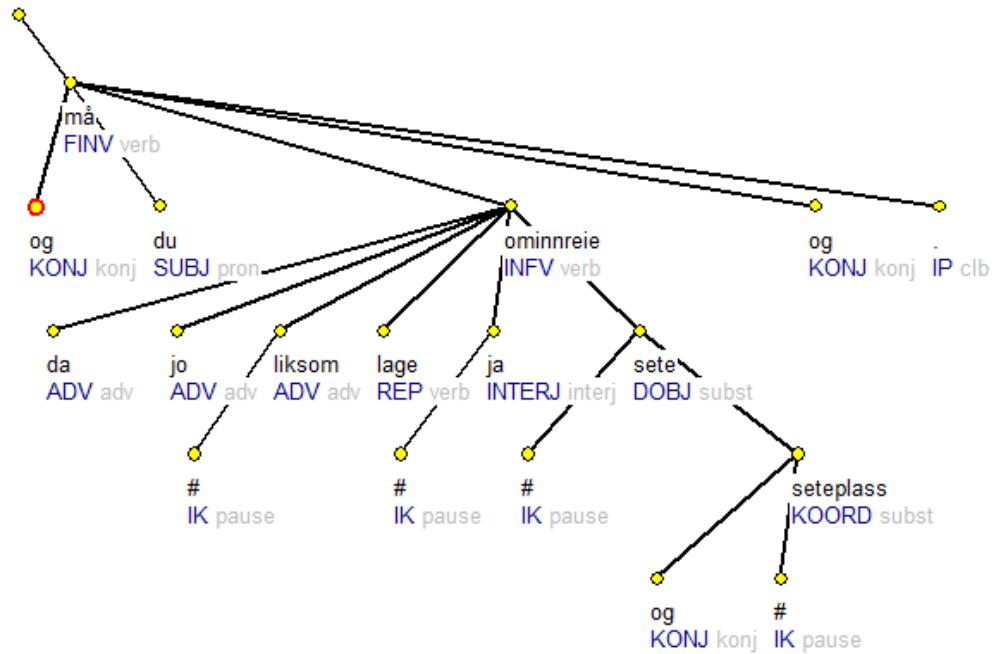
Det er vanlig å avslutte en ytring i talemål med en konjunksjon (og, men, for, eller og så). Det er så å si umulig å vurdere hva slags funksjon disse skal ha for vi vet ikke hva informanten hadde tenkt å fortsette å si. Ofte vil disse fungere mer som en diskursartikkelen som viser at ytringen er ferdig. Vi har bestemt at disse skal ha funksjonen KONJ og henge på rotnoden.

(53) får nå den avvekslinga at du har nå vore og fiska **så**. (Austevoll_uib_01 ... 0005 2/50)

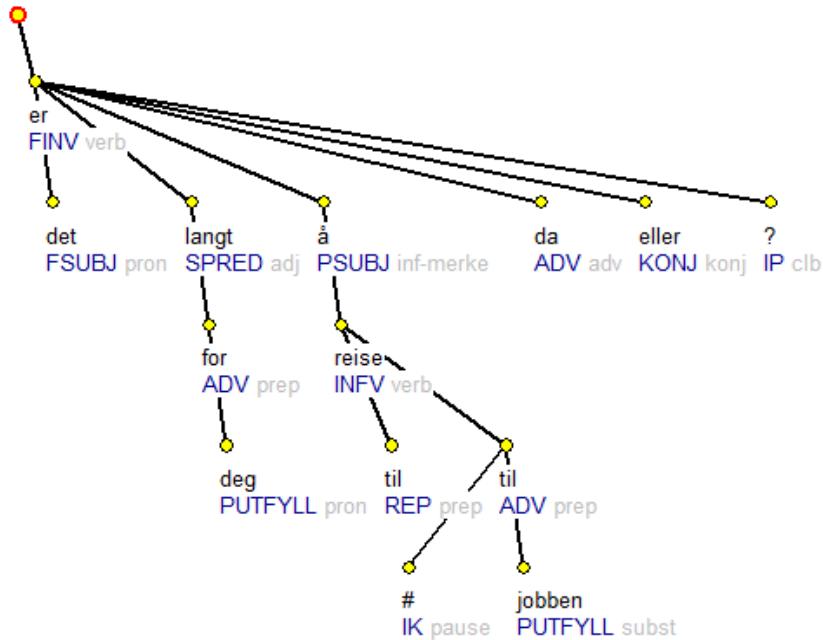
(54) for mykje om # eller for mykje og for mykje det er jo klart at vi er # ute å køyre **men**.

(Austevoll_uib_01 ... 0005 34/50)

(55) og da må du jo # liksom lage # ja ominnreie # sete og # seteplass **og** . (austevoll_uib_01 ...
0002 13/50)



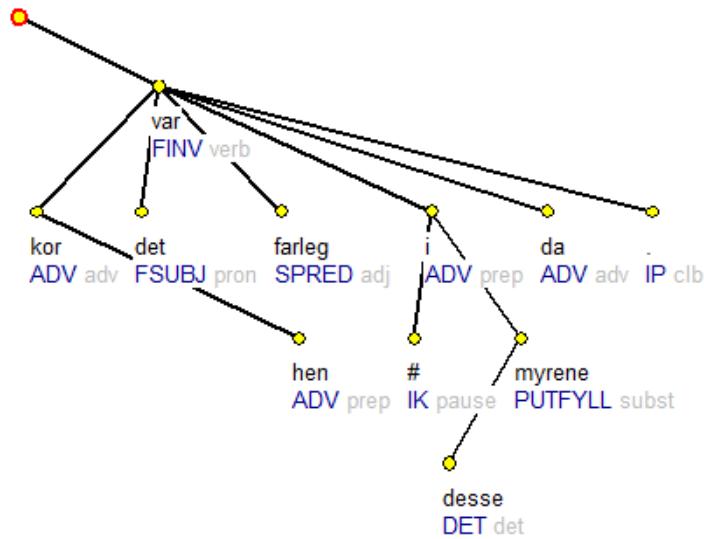
(56) er det langt for deg å reise til # til jobben da **eller**? (austevoll_uib_01 ... 0002 33/50)



12.3 Da på slutten av segment

Ofte kan et segment avsluttes med *da*. Dette er ikke av typen setningsadverbial, men mer som en diskurspartikkel som viser at ytringen er avsluttet. Den skal da annoteres som ADV og henges på rotnoden, parallelt med konjunksjoner til slutt i segmentet (se avsnitt 12.2).

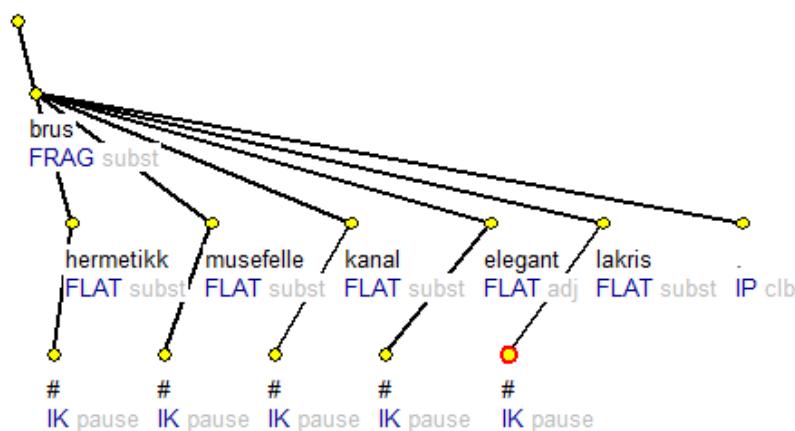
(57) kor det var farleg hen # i desse myrene da . (aamot_uio_01 ... 0000 21/423)



12.4 Oppramsing, opplesning av lister

Informanten ramser opp ord fra ei liste i fila. I sånne tilfeller bare setter vi første ord som rot (FRAG), og resten av ordene som FLAT. Eksempelet i (58) viser et segment som består av opplesing av en liste med ord.

(58) brus # hermetikk # musefelle # kanal # elegant # lakris (trondheim_uib_10 ... 0000 160/216)

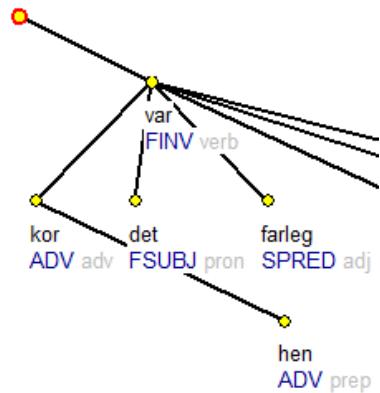


13 Frekvente «faste» uttrykk.

Dette er en del av de frekvente uttrykkene som man fort kan bli usikker på hvordan skal analyseres.

Denne delen kan gjerne utvides fortløpende når man kommer over faste uttrykk.

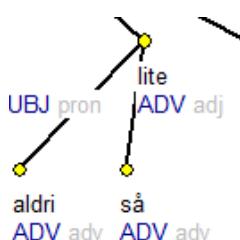
(59) kor hen



(60) lite grann



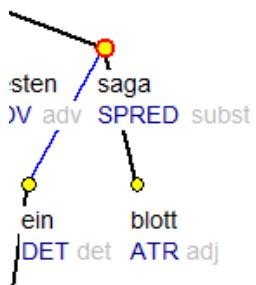
(61) aldri så lite



(62) om å gjere



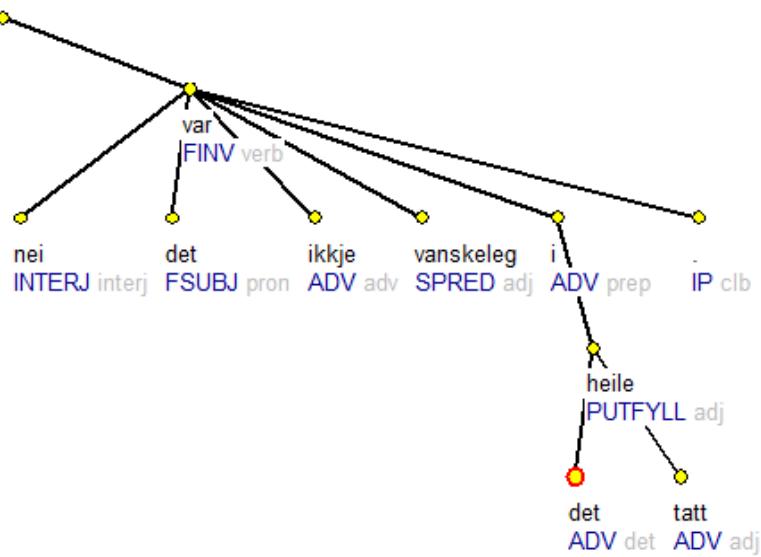
(63) en saga blott



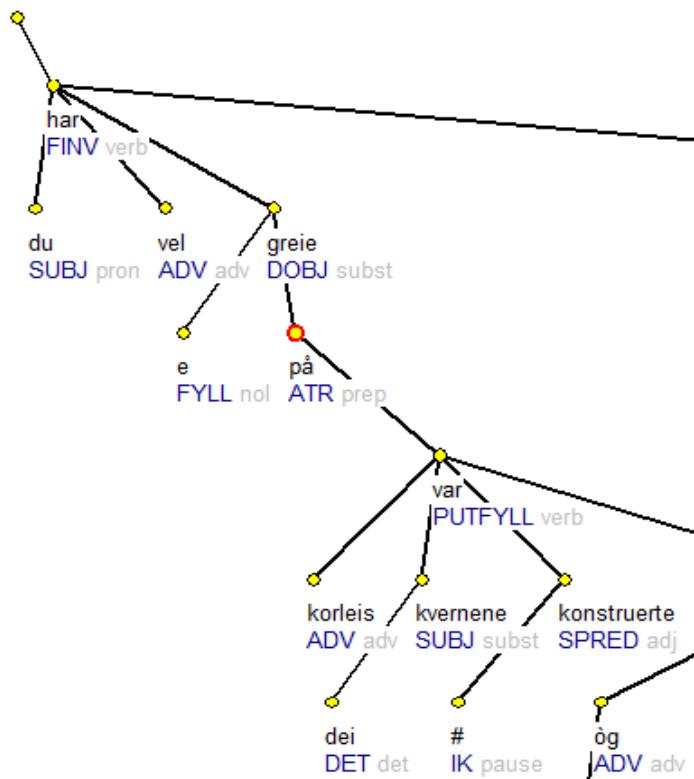
(64) å så reise (splittet infinitiv)



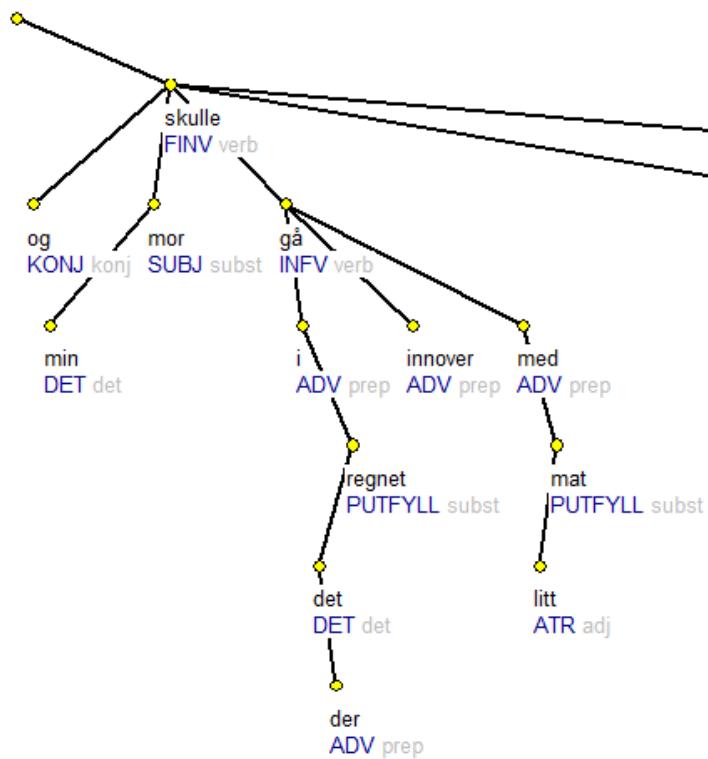
(65) I det hele tatt



(66) Ha greie på



(67) Det der



Merk at «der» kan fungere som ADV eller ATR, avhengig av hodets ordklasse. Denne analysen gjelder kun i tilfeller hvor «der» utpeker eller identifiserer noe.