

NDT Guidelines for Morphological and Syntactic Annotation

Kari Kinn, Per Erik Solberg og Pål Kristian Eriksen
Translated from Norwegian to English by Per Erik Solberg

Contents

I	Morphology	6
1	Morphological annotation	7
1.1	Morphological analyses from OBT	7
1.1.1	Nouns with optional gender	7
1.1.2	The distinction between preposition and complementizers	7
1.1.3	The <i>clb</i> tag	8
1.1.4	Tags with angle brackets	8
1.1.5	Classifying prepositions	9
1.1.6	Proper Names	9
1.1.7	Heading token	9
1.2	Analyses which do not follow OBT	9
1.2.1	Participles	10
1.2.1.1	Perfect participle in Bokmål	10
1.2.1.2	Perfect participle in Nynorsk	10
1.2.2	<i>Siden</i>	11
1.2.3	Non-standard	11
1.2.3.1	Foreign language and dialect in the text	11
1.2.3.2	Non-standard word division	12
1.2.3.3	Typos	13
1.2.3.4	Non-standard inflection	13
1.2.3.5	Lack of congruence	14
1.2.3.6	Lack of congruence between plural subjects and participles in Nynorsk	14
1.2.4	Coordination of compounds where a repeated element is omitted	15
1.2.4.1	Not proper nouns	15
1.2.4.2	Proper nouns	15
1.2.5	Tokenization	15
1.2.5.1	Whitespace	16
1.2.5.2	Quotation marks and parentheses	16
1.2.5.3	Punctuation	16
1.2.6	Morphological analysis of complex expressions	16
1.2.6.1	Complex expressions with whitespace	16
1.2.6.2	<i>Au pair, hokus pokus</i> etc.	16
1.2.6.3	Idiomatic PPs	17
1.2.7	<i>Selv</i>	17
1.2.8	Floating quantifiers	19
1.3	Sentence division	19

II	Syntax	21
2	Introduction	22
2.1	Dependency Grammar	22
2.2	Fundamental principles of analysis	23
2.2.1	No empty nodes	24
2.2.2	Only unique dependencies	24
2.2.3	Crossing dependencies	24
2.2.4	Lexical words and function words	26
2.3	Overview of the syntactic tagset	26
3	Sentences	28
3.1	Root functions	28
3.1.1	FINV	28
3.1.2	INTERJ	29
3.1.3	FRAG	30
3.1.4	The relationship between auxiliary verbs and lexical verbs	31
3.2	Subjects	33
3.2.1	Referential subjects in subject position (SUBJ)	33
3.2.2	Formal subjects (FSUBJ)	34
3.2.2.1	Weather sentences and related constructions	34
3.2.2.2	Presentational sentences	35
3.2.2.3	Impersonal passive sentences	35
3.2.2.4	In clause-anticipating constructions:	35
3.2.2.5	Clefting	36
3.2.3	Potential subjects (PSUBJ)	36
3.2.4	Formal and potential <i>objects</i> (FOBJ and POBJ)	38
3.2.5	How to distinguish between potential subjects/objects and adverbials	39
3.2.6	How to distinguish between potential subjects and subject predicatives	41
3.3	Clefting	43
3.4	Objects	47
3.4.1	Direct objects (DOBJ)	47
3.4.1.1	Complements to verbs	47
3.4.1.2	Complements to adjectives	48
3.4.2	Indirect Objects (IOBJ)	49
3.4.3	Perception verbs and <i>be, la</i> etc. with object and infinitive	50
3.4.3.1	Perception verbs: “objects” as subjects for the infinitive	50
3.4.3.2	<i>Be, la</i> etc.: IOBJ + DOBJ	51
3.5	Adverbials (ADV)	53
3.5.1	ADV to verbs	54
3.5.2	ADV to non-verbal heads: adjectives, determiners, ad- verbs and prepositions	55
3.5.3	ADV to nouns	62
3.5.4	Complex time adverbials	62
3.5.5	Path-constructions	62
3.6	How to distinguish between transitive and intransitive prepositions	64
3.7	Non-finite verbs	66

3.7.1	Auxiliary verbs and lexical verbs	67
3.7.2	Infinitival clauses with an infinitival marker	67
3.7.3	Infinitival clauses without an infinitive marker	70
3.7.4	Participles	70
3.8	Predicative	72
3.8.1	Copular verbs	73
3.8.2	OPRED with <i>ha</i>	74
3.8.3	Predicative after <i>med</i>	75
3.8.4	FOPRED to non-objects	77
3.9	Interjections	77
3.10	Subordinate clauses	79
3.10.1	Fundamental choices of analysis	79
3.10.2	The complementizer	80
3.10.3	Relative clauses	82
3.10.4	Indirect interrogative clauses etc.	86
4	Dependents with non-verbal heads	89
4.1	The nominal phrase	89
4.1.1	Introductory remarks	89
4.1.2	ATR	91
4.1.2.1	Different types of descriptors	91
4.1.2.2	<i>Av</i> -partitive	93
4.1.3	DET	94
4.1.3.1	Genitive	94
4.1.3.2	Quantity nouns	95
4.1.3.3	Quantity designations with a preposition	99
4.1.4	Cardinal numbers following a noun	100
4.2	APP	102
4.2.1	APP in nominal phrases	103
4.2.2	Special rules for APP with proper nouns	105
4.2.3	APP when dislocated positions are duplicated	109
4.2.4	APP with <i>så</i>	111
4.2.5	APP with <i>noe</i>	112
4.2.6	Appositions with the abbreviation <i>dvs.</i>	113
4.2.7	Subset appositions	114
4.3	Dependents on prepositions	119
4.3.1	Prepositional complement	120
4.3.2	Postponed prepositions	122
4.3.3	ADV on prepositions	122
4.3.4	Predicative with <i>med</i>	123
4.3.5	New prepositions	123
4.4	Coordination	124
4.4.1	Coordination with two conjunctions	128
4.4.2	Coordination without conjuncts	128
4.4.3	<i>Enten-eller</i> coordination as adverbial clauses	130
4.4.4	Adverbs and adverbial phrases as quasi-conjunctions	132
4.5	Flat structure	134
4.5.1	Quotes in foreign languages	134
4.5.2	Complex expressions with whitespace, <i>au pair</i> og non-standard word division	135

4.5.3	Proper nouns	136
4.5.4	Numbers with whitespaces	136
4.5.5	Complex words with a slash	137
5	Other questions	138
5.1	Ellipsis	138
5.1.1	Noun ellipsis	138
5.1.2	Elliptical matrix clauses	142
5.1.3	Ellipsis and coordination	143
5.1.3.1	Coordination with noun ellipsis	144
5.1.3.2	Coordination with verb ellipsis	145
5.1.4	Ellipsis and the prepositions <i>som</i> and <i>enn</i>	147
5.2	Punctuation	149
5.2.1	Sentence punctuation	149
5.2.2	Sentence-internal punctuation	150
5.2.3	Comma	151
5.2.4	Quotation marks and parentheses	154
5.3	Parenthetical expressions	155
5.3.1	Address	156
5.3.2	Parenthetic clauses	157
5.3.3	Utterance verbs in direct speech	158
5.3.4	Parenthetic verbs in the imperative	159
5.3.5	Utterance verbs at the beginning of the sentence	159
5.4	Superfluous words in the sentence - UKJENT	160
5.5	<i>Siden</i>	161
5.5.1	<i>Siden</i> as an adverb	161
5.5.2	<i>Siden</i> as a preposition	162
5.5.3	<i>Siden</i> as a complementizer	162
5.6	Quantity indications of the type <i>to og et halvt år</i>	163
5.7	<i>Ha vanskelig for</i>	163
5.8	Comparative constructions	164
5.8.1	The head in comparative constructions	164
5.8.2	<i>Slik/samme som</i>	171
5.8.3	<i>Jo mer, jo bedre</i>	173
5.9	Mixed quotation	174
5.10	Garpe-genitive	176
5.11	Section references etc.	178
5.12	Tough movement	179
5.13	Fragmentary wh-questions	181
5.14	Fragmentary conditional clauses	181
5.15	Multiple adverbials in the left periphery	182
5.16	<i>Selv</i>	183
5.17	Floating quantifiers	185
5.18	Difficult sentences	186

Introduction

This document presents the annotational choices in the Norwegian Dependency Treebank (NDT), a treebank for Norwegian Bokmål and Nynorsk developed by Språkbanken at the National Library of Norway from 2011 to 2013. These guidelines are written by the annotators in treebanks. We have had two audiences in mind: Firstly, these guidelines have been used as reference during annotation. Secondly, they are intended as explanations of our analyses for users of the treebank. We largely follow the Norwegian Reference Grammar (NRG), (Faarlund et al. 1997), and we refer to NRG for a detailed description of the structures discussed in this document. When there are large discrepancies between our analysis and those in NRG, this will be explicitly mentioned. The annotations in the treebank are also based on the more general principles set up in Johannessen et al. 2011, a report written at the beginning of the project period.

The majority of the choices documented here have been made in parallel with the annotation of the first texts in the corpus, during the first six months of the project, and the first edition of this document was written towards the end of this initial phase. Since then, a number of the initial analytical choices have been changed, and the annotation of constructions we hadn't thought of in the initial phase, have been documented. We have tried, whenever possible, to illustrate our analyses with pictures of dependency trees from the corpus. When we have documented updated annotations after the first edition of the guidelines, we have obviously also updated the image which illustrates the annotation. However, we have not given priority to reviewing all the illustrations which appear in this document for each new edition. There may therefore be some examples of old analyses in some of the pictures, but the analysis of the phenomenon of image is meant to illustrate, will always be correct.

The document is divided into two main parts. Part I presents the choices we have made in the morphological analysis. This part is relatively short, since we mostly follow the *Oslo-Bergen Tagger* and have made only a few independent choices in the morphological analysis. In part II we present the syntactic annotation. Chapter 2 gives a brief introduction to Dependency Grammar and accounts for the general principles of annotation. In chapter 3 we document the annotation of constructions which belong in the clausal domain, while chapter 4 deals with the analysis of noun phrases, prepositional phrases and coordination. In chapter 5 we discuss some analysis and issues that have not been addressed in the previous chapters.

Part I

Morphology

Chapter 1

Morphological annotation

Introduction

The texts in NDT are tagged with *the Oslo-Bergen Tagger* (OBT; Lynum et al. 2011), and then manually disambiguated. We follow the morphological analyses in OBT to a large extent, but make some independent choices. We will not describe OBT’s analyses in detail, but refer to the chapters on morphology in Faarlund et al. 1997, which OBT follows closely, and the documentation on OBT’s web page. We found it necessary to comment on a few of the analyses from OBT. This is done in section 1.1. 1.2 describes the cases where we have made made different choices from OBT. 1.3 describes the sentence division in the corpus.

1.1 Morphological analyses from OBT

1.1.1 Nouns with optional gender

Many nouns in Bokmål can be either masculine or feminine. When it is not possible to determine based on morphology or syntax which gender the noun has, feminine is chosen. Some nouns in Nynorsk also have optional gender, such as *bluse*, ‘blouse’, and *kasse*, ‘box’. Feminine is chosen also in these cases.

1.1.2 The distinction between preposition and complementizers

Some words can either be prepositions and complementizers, such as *som*, ‘that, as’, *enn*, ‘than’, and *etter*, ‘after’. When such a word is followed by a clause, it is marked as complementizer if the clause has no other complementizer, and preposition if the clause has another complementizer. In (1.1 a) *enn* shall be marked as complementizer, but in (1.1 b) *enn* is a preposition, as it is followed by the complementizer *at*.

- (1.1) (a) *Men virkeligheten er annerledes enn jeg trodde da jeg*
but reality+the is different than I believed when I
var barn.
was child

'But the reality is different from what I believed when I was a child.' (Dagbladet 003)

- (b) *Alternativa er verre enn at Hellas framleis er med.*
alternatives+the are worse than that Greece still is
with

'The alternative is worse than Greece still being a part of it.' (Dag og Tid 002)

1.1.3 The *clb* tag

OBT is a rule-based tags using Constraint Grammar (CG). CG rules often refer to clause boundaries. Wh-words and complementizers are unambiguous clause boundaries and need not be marked in any particular way. Conjunctions, however, may, but need not, mark a sentence boundary. In (1.2 a) the conjunction marks a clause boundary, while it does not in (1.2 b):

- (1.2) (a) *Per kjøper epler og Kari spiser is.*
Per buys apples and Kari eats ice cream
'Per buys apples and Kari eats ice cream.'
- (b) *Per og Kari kjøper epler.*
Per and Kari buy apples
'Per and Kari buy apples.'

In order to have unambiguous clause boundaries also with conjunctions, clause-separating conjunctions have the tag *clb* (*clause boundary*). Note that when another clause-separating element comes after the conjunction, *clb* shall not be used, as the clause boundary is marked by this second element. In (1.3), the conjunction shall not have a *clb*-tag, as it is followed by a wh-word.

- (1.3) *Hva heter mannen din og hvor kommer han fra?*
what is called man+the yours and where comes he from
'What is the name of your husband and where does he come from?'

Conjunctions which do not mark clause boundaries, have the tags *<ikke-clb>* (Bokmål) or *<ikkje-clb>* (Nynorsk). We initially wanted commas to be marked as either *clb* or *<ikke-clb>* too, but we had to drop this for technical reasons.

1.1.4 Tags with angle brackets

Some words have certain morphological tags in angle brackets. These tags provide additional information about the word which may be relevant, but does not concern the morphological analysis directly. In (1.4 a), the adjective that adjective *grillet*, 'grilled', is formed on the basis of the perfect participle of the verb *grille*, 'grill', and therefore has the tag *<perf-part>*. In (1.4 b), the tag *<adv>* indicates that *lenge*, 'long', previously was classified as an adverb.

- (1.4) (a) "grillet" "grille" adj nøyt ub ent *<perf-part>*
(b) "lenge" "lenge" adj pos *<adv>*

Users of the corpus can choose to remove these tags, if they wish. Note, however, that some punctuation tokens have a tag in angle brackets as part-of-speech tag and only morphological tag:

(1.5) (a) “\$,,” <komma>

1.1.5 Classifying prepositions

The class of prepositions is somewhat larger in Faarlund et al. (1997), and consequently in OBT, than what has been customary in traditional grammar instruction. For example, *her*, ‘here’, *der*, ‘there’, *hit*, ‘over here’ and *dit*, ‘over there’, are classified as prepositions. The same is true for *nord*, ‘north’, *sør*, ‘south’, *øst*, ‘east’, and *vest*, ‘west’. See also 4.3.

1.1.6 Proper Names

OBT consistently tags words beginning with capital letters as proper names (postag: subst, features: prop), except at the beginning of a sentence or in compounds such as *Ibsen-forestilling*, ‘Ibsen play’, a practice we follow. Uncapitalized words can sometimes be tagged as nouns, e.g. names of some companies, bands, organizations etc.: *finn.no*, *a-ha*.

In the case of complex proper names, such as *Den norske kirke*, *Universitetet i Oslo* and *Næringslivets hovedorganisasjon*, the capitalized words shall be tagged as proper names, while the rest shall have their usual morphological tags.

One of the few morphological features OBT allows for proper names, is the genitive tag. This tag is used when the proper name functions as a possessor or is a part of a phrase with this function, as in *Pers bil* or *broren til Pers bil* (cf. 4.1.3.1). When the first word in a complex proper name is a genitive, as in *Næringslivets hovedorganisasjon*, the genitive tag shall not be used. For foreign names, see 1.2.3.1. For the syntactic analysis of complex proper names, see 4.5.

1.1.7 Heading token

After headlines etc. which do not contain sentence-final punctuation, OBT puts in a new token with the word form |, lemma \$|, part of speech *clb*, and the morphological tag <*overskrift*> (‘heading’). We have chosen to retain such tokens in the corpus, mainly for technical reasons. Heading tokens can be deleted by users of the corpus without the other analyses being harmed, as this token is always sentence-final, is always dependent on the head of the sentence, and never has its own dependents.

1.2 Analyses which do not follow OBT

Here we present the cases where the morphological analyses in NDT differ from OBT’s analyses, or where the choices made in NDT do not follow directly from those made in OBT.

1.2.1 Participles

1.2.1.1 Perfect participle in Bokmål

In Bokmål, OBT gives both *adjective* and *verb* as possible parts of speech for participles. In both cases, the lemma is the infinitive of the verb from which the participle is derived. Adjective participles have the morphological tag *<perf-part>*. We follow these guidelines when we choose part of speech for adjectives:

When the participle is an attribute which precedes a head noun, the part of speech shall be *adjective*. If the participle is a complement to the auxiliary verbs *ha*, *få*, or to a modal, the part of speech should be *verb*¹. When the participle is a complement to the verbs *være*, 'be', and *bli*, 'become', we choose either *verb* or *adjective*. (If we choose verb, the syntactic function should be INFV. If we choose *adjective*, the function should be SPRED.) We use the following test to choose between parts of speech in such cases: If you can insert a manner adverbial such as *fort*, 'quick', *sakte*, 'slow', etc., the part of speech should be *verb*. If you can insert a degree adverbial, such as *veldig*, 'very', the part of speech should be *adjective*. When the participle is an attribute following its head noun, as in (1.6 a) or a predicative (free or obligatory), as in (1.6 b), we must also choose between *verb* and *adjective* using the same test. In (1.6 a), we have chosen the part of speech *verb*, as you can modify *født* with a manner adverbial such as *med møyse*, 'with difficulty'. In (1.6 b), *veldig* can modify *nedbrutt*. The part of speech is therefore *adjective*.

- (1.6) (a) *En fattig jente født i et fattig land*
 a poor girl born in a poor country
 'A poor girl born in a poor country.' (Aftenposten 001)
- (b) *Han tok nedbrutt imot den triste nyheten.*
 He took broken particle the sad news
 'Broken, he recieved the sad news.'

For lexicalized adjectives formed on the basis of a participle, such as *interessert*, 'interested', OBT gives three options: *verb*, *adjective* with the feature *<perf-part>* and infinitival lemma, and *adjective* with an adjectival lemma and no *<perf-part>* feature. In such cases, we either choose *verb* or the lexicalized adjective form, i.e. the one that has an adjectival lemma and no *<perf-part>* feature.

1.2.1.2 Perfect participle in Nynorsk

In Nynorsk, the situation is more complicated: Participles of e-verbs and strong verbs are congruent with the subject after the verbs *vere*, 'be' and *verte/bli*, 'become', regardless of whether these verbs function as auxiliaries or not, while a-verbs do not have visible congruence. After the auxiliaries *ha*, *få*, and modal verbs, the neuter form of the participle is always used. OBT will give a choice between *adjective* and *verb* for neuter participles, but not for masculine or feminine participles. In other words, for many verbs we will not have the opportunity to choose between *adjective* and *verb*, as we do in Bokmål. For Nynorsk, we follow to the following procedure:

¹When the verb *få* is used in constructions such as *få maten servert*, 'get food+the served'. We choose *adjective*, and give the participle the syntactic function OPRED (*object predicative*) cf. 3.8.

1. After *ha* and *få*, used as auxiliaries, and modal verbs, we always choose *verb*, as in Bokmål.
2. In all other cases, we choose *adjective*, regardless of whether OBT allows a verb reading.
3. In the case of lexicalized participles, we choose the reading with the *<perf-part>* tag in the cases where we would have chosen *verb* in Bokmål, and a reading without *<perf-part>* in the cases where we would have chosen *adjective* in Bokmål.

Participles after *vere* and *verte/bli* will have the syntactic functions INFV and SPRED. It will therefore be possible to change the part of speech to *verb* for all participles with the function INFV for users who are interested in this. This will not be possible for predicatives which are not dependent on *vere* and *verte/bli* and for attributes which follow after a head noun, however. See 3.7.4 for the syntactic analysis of participles.

1.2.2 *Siden*

In OBT, *siden*, 'since' have *adverb* and *complementizer* as possible parts of speech. We believe *siden* is a preposition in cases like (1.7), where it is followed by a complement (cf. Faarlund et al. 1997, 413). We therefore manually add the part of speech *preposition* in such cases.

- (1.7) *Etter å ha jobbet seks og syv dager i uka*
 after to have worked six and seven days in week+the
siden 1992, har jeg endelig den friheten og kontrollen
 since 1992 have I finally the freedom+the and control+the
jeg har drømt om.
 I have dreamt about
 'After having worked six and seven days a week since 1992, I finally have the freedom and the control I have dreamt of.' (Dagbladet 001)

See also 5.5.

1.2.3 Non-standard

Sometimes we find non-standard words and phrases in the text. They can be of different types:

1.2.3.1 Foreign language and dialect in the text

Words in foreign languages or dialects used in the Norwegian text should be tagged with part of speech, lemma and any other relevant morphological tags, as far as possible. In addition, they have the tag *unorm* (non-standard). If it is not possible to determine the part of speech, we set it to *ukjent*, 'unknown' (without the *unorm* tag). In (1.8 a), *eg*, *e*, and *sjenerte* must have an *unorm* tag, as these words are non-standard in Bokmål. The lemma should be the same as for the standard Bokmål lemma. In (1.8 b), *bad* should have *adjective* as part of speech and the *unorm* tag. *For* and *business* are standard Bokmål forms, and should therefore not have any special tags. For words such as *bad*,

for which no standard spelling variant exists, the lemma should be the same as the word form.

(1.8) (a)

Eg e ganske sjenerte.
I am relatively shy
'I am relatively shy' (Dagbladet 003)

(b) *Hvis vi forstyrrer en kunde, er det bad for business,*
if we disturb a customer is that bad for business
sier Kristensen.
says Kristensen
'It is bad for business if we disturb a customer, says Kristensen.'
(Dagbladet 004)

If the text reproduces a quotation in a foreign language, as in (1.9), we set *ukjent* as part of speech of all words in the quote. Also the word with a capital letter shall have *ukjent* (unlike complex organization names etc., see below), while punctuation, quotation marks, etc. shall have their usual morphological tags. In syntax, such quotes have a flat structure, see 4.5.

(1.9) *I de internasjonale reglene står det blant annet at*
in the international rules+the stands it among other that
“the buttocks must be covered at all times”.
eng eng eng eng eng eng eng eng
'In the international rules, it says, among other things, that “the
buttocks must be covered at all times”.(Dagbladet 002)

It is often not clear whether a sequence of foreign words is to be regarded as a quote or as part of the Norwegian text. This is therefore something the annotators must decide in each case.

In the case of complex organization names etc. in foreign languages, where all words have capitalized initials, such as *London City Airport*, all words should be tagged as proper nouns. This also applies if one or more of the words are not nouns in the language, as in *Take That*. If the complex foreign name contains a word or words with lowercase first letter, such as *United States of America*, the non-capitalized words should receive *unorm* as part of speech.

See also 4.5.

1.2.3.2 Non-standard word division

When the parts of a compound are separated by a whitespace, as in (1.10), which is non-standard in Norwegian, we do not tokenize the parts into one single token, but we tokenize them as separate tokens and give them an adequate morphological analysis to the extent that it is possible. When the word forms are not found in the dictionary, we use the *unorm* tag.

(1.10) *Din favoritt roman er?*
your favorite novel is
'Your favorite novel is?' (Dagbladet 003)

See also 4.5.

1.2.3.3 Typos

In the case of typos, such as *prostiuerte* for *prostituerte*, ‘prostitutes’, in (1.11 a), we add correct morphological tags as far as possible, and add the *unorm* tag. Lemma will be the normalized lemma form. In the case of misspelled words found in the dictionary, but not with the intended meaning, as in (1.11 b), we use the morphological tags and lemma of the intended word, not the homophonous one. In addition, we add the *unorm* tag. *En* should therefore not be the determiner *en*, ‘one’, but a preposition with the lemma *enn*, ‘than’.

- (1.11) (a) *Jeg er svært sjokkert over at noen kan lage et*
 I am very shocked over that someone can make a
slikt teaterstykke, sier den tidligere prostiuerte Lita
 such play says the former prostitute.misspelled Lita
Malmberg til Politiken.
 Malmberg to Politiken
 ‘I am very chocked that someone can make such a play, the former
 prostitute Lita Malmberg says to Politiken.’ (Dagbladet 004)
- (b) *Idrettsgrenen, som på norsk kalles*
 athletic discipline+the which on Norwegian is called
sportsdans, har skapt en klessjanger som går i en
 sport dance has created a clothing genre which goes in a
helt annen takt en motebildet for
 whole other rythm than.misspelled fasion picture+the for
øvrig [...].
 general
 ‘The athletic discipline, which in Norwegian is called “sportsdans”,
 has created a clothing style which develops at a different pace than
 the fasion scene in general.’ (Dagbladet 002)

A spelling is considered non-standard if it is not found in the lexical database Norsk ordbank or is a compound. A consequence of this is that the words which are acceptable in Norwegian, but are not found in Norsk ordbank, such as *haugesundsk* (Dagbladet 003), ‘in Haugesund dialect’, must have an *unorm* tag. Compounds shall have an *unorm* tag if the last element of it is not in Norsk ordbank. Other elements of the compound will not be taken into consideration when assigning the *unorm* tag. *Kyllingfillet* (blogg-bm 001), ‘chicken filet’, is marked with *unorm*, because the last term, *fillet*, is not a spelling which is found in Norsk ordbank. *cupcakeinnspill* (blogg-bm 001), ‘cupcake suggestion’, is not marked with *unorm*, however, as only the first element, *cupcake*, is lacking in Norsk ordbank. This is a consequence of how OBT treats compounds, and we have chosen to follow it, to get a tagging which is as consistent as possible.

1.2.3.4 Non-standard inflection

In (1.12) used in *dogg*, ‘dew’, is neuter, which is not standardized Bokmål. The word should have a *neuter* tag and an *unorm* tag:

- (1.12) *Vindusviskerne stryker som fiolinbuer over*
 windshield wipers+the stroke as violin bows over
vinduene, og Staveland titter ut gjennom dogget.
 windows+the and Staveland peeks out through dew+the.neut

'The windshield wipers stroke as violin bows over the windows, and Staveland peeks out through the dew.' (Dagbladet 003)

1.2.3.5 Lack of congruence

When two words are not congruent, but both are standardized forms, we do not use the *unorm* tag. In (1.13) is *påle*, 'pole', preceded by a feminine determiner. *Påle* cannot be feminine in standard Norwegian, only masculine. The determiner will have a *feminine* tag and the noun a *masculine* tag, but non of them have an *unorm* tag.

- (1.13) *Vi er ikke her for å sette ei påle i jorda og*
 we are not here for to place a.fem pole in earth+the and
si at Europa er vårt.
 say that Europe is ours
 'We are not here to place a pole in the earth and say that Europe is ours.' (Dagbladet 001)

1.2.3.6 Lack of congruence between plural subjects and participles in Nynorsk

In Nynorsk, there is congruence of number between subjects and participles which are complements to the verbs *vere*, 'be', and *verte/bli*, 'become'. There are frequent examples of lack of such congruence, however. When there is a number mismatch between a plural subject and participle in Norwegian, and the singular form of the participle is ambiguous between neuter and masculine/feminine, as for *høyrð* in (1.14), we mark the participle as neuter.

- (1.14) *Elevane vart ikkje høyrð.*
 pupils were not heard.sing
 'The pupils were not heard.'

It is probably descriptively adequate to say that language users choose a default neuter form in such examples in Nynorsk, as they do in Bokmål. (The alternative is to say that we have congruence for gender, but not number, but that seems less likely.)

In other cases of lack of congruence of number, such as predicative adjectives and attributive participles following the head ((1.15 a) and (1.15 b)), we assume congruence of gender:

- (1.15) (a) *Eg veit at dei fleste foreldra i Fyresdal er*
 I know that the most parents+the in Fyresdal are
redd utviklinga...
 afraid.sing development+the
 'I know that most parents in Fyresdal are afraid of the development'
 (Vest-Telemark Blad 005)
- (b) *Det var óg tilbod om omvisning til stadene*
 there was also offer about guided tour to places+the
knytt til Storegut og Dyre Vaa-samlingane.
 tied.sing to Storegut and Dyre Vaa collections+the
 'A guided tour of the places linked to Storegut and the Dyre Vaa collections was also offered.' (Vest-Telemark Blad 005)

1.2.4 Coordination of compounds where a repeated element is omitted

In the case of coordination of compounds where one of the elements is repeated, it is normal to omit the repeated element in one or more of the conjuncts and represent the omitted element with a hyphen: *regjerings- og opposisjonssiden*, 'government and opposition side', *lærerutdanninga og -yrket*, 'teacher's education and profession', *Arbeids- og velferdsetaten*, 'Labour and Welfare Agency'. The analysis depends on whether the coordination phrase begins with a capital letter, and therefore should be considered as a proper noun, or not.

1.2.4.1 Not proper nouns

The main principle for this kind of coordination of compounds which are not proper names, is that the conjuncts with the omitted element shall have the same part-of-speech tag as that last conjunct and the morphological tag *ufl* (*ufullstendig*, 'incomplete') if it is the last element which is omitted, as it is the last element which is essential for part of speech and morphology. The lemma is the same as the word form for all words marked with *ufl*. If it is the first element which is missing, the words with the missing element will have complete morphological tags, in addition to *samset*, the tag which marks compounds. The analysis of *regjerings- og opposisjonssiden* is therefore as follows, since it is the last element which is omitted.

- (1.16) *form*: regjerings- *lemma*: regjerings- *pos*: subst *features*: ufl
form: og *lemma*: og *pos*: konj *features*: <ikke-clb>
form: opposisjonssiden *lemma*: opposisjonsside *pos*: subst *features*:
 appell|mask|be|ent|samset

In *lærerutdanninga og -yrket*, the first element is omitted in the last conjunct. The morphological analysis is therefore as follows:

- (1.17) *form*: lærerutdanninga *lemma*: lærerutdanning *pos*: subst *features*:
 appell|fem|be|ent|samset
form: og *lemma*: og *pos*: konj *features*: <ikke-clb>
form: -yrket *lemma*: -yrke *pos*: subst *features*: appell|nøyt|be|ent|samset

Note that the hyphen is included in the lemma for *-yrket*, as it represents the first element of the compound.

The syntactic analysis is the same as for regular coordination, cf. 4.4.

1.2.4.2 Proper nouns

If the first conjunct in such a coordination phrase has a capital letter, as in *Arbeids- og velferdsetaten*, we consider the phrase as a complex proper noun, and analysis shall be as other proper nouns (cf. 1.1.6): *Arbeids-* shall be tagged as a proper noun, the other words have their usual morphological tags.

1.2.5 Tokenization

The tokenization in NDT differs somewhat from the one chosen in OBT. In the following, we explain the principles we have followed and point out when they deviate from OBT's.

1.2.5.1 Whitespace

We always split tokens on whitespaces in NDT. This is not always the case in OBT, where some fixed expressions, as *i går*, 'yesterday' and *au pair*, are tokenized as one token. For the morphological and syntactic analysis of such fixed expressions, see 1.2.6 and 4.5.

1.2.5.2 Quotation marks and parentheses

If quotation marks stand on both sides of a single token, *Han kaller seg "konge"*, 'he calls himself "king"', or if the quotes are inside a token, *"er det ikke slik at"-spørsmål*, "'isn't it so" questions', OBT does not make the quotation marks separate tokens. Quotation marks will always be separate tokens in NDT, however, also in examples such as these. In *"er det ikke slik at"-spørsmål*, "*er det ikke slik at*" and *-spørsmål* are therefore separate tokens. The same is true for parentheses in similar environments.

1.2.5.3 Punctuation

A comma is always a separate token, except in numbers that contain commas, such as *0,6*. Semicolons, exclamation marks and question marks are always separate tokens. A colon is a separate token, except in numbers that contain colons, *10:13*, and in URLs, *http://www.nb.no*.

Full stops are usually separate tokens. Exceptions are numbers that contain a full stop, *1.3*, abbreviations, *f.eks.* and *Carl J. Hambro*, and URLs, *www.nb.no*. Similarly, a slash is its own token, except in URLs, *http://www.nb.no*, and in some common abbreviations, like *A/S*.

1.2.6 Morphological analysis of complex expressions

Since a whitespace always leads to a token split in NDT, cf. 1.2.5.1, the individual parts of complex expressions must get an independent morphological analysis. This will be described here.

1.2.6.1 Complex expressions with whitespace

Expressions such as *New York Rangers-fanklubben*, 'New York Rangers fan club', functions, syntactically speaking, as one word, even it contains whitespaces. Since we do not accept tokens with whitespaces, such expressions must be split up. All words have their regular morphological tags: *New* and *York* are tagged as proper nouns, *Rangers-fanklubben* as a compound noun. We show that such expressions constitute a syntactic unit by giving them a flat structure in the syntactic analysis, see 4.5.

1.2.6.2 *Au pair*, *hokus pokus* etc.

Some phrases which are borrowed from other languages, such as *au pair* and *hokus pokus*, 'hocus-pocus', function in Norwegian as one word, and the individual parts of the phrase do not have any meaning on its own. In such cases, all parts of the expression have the same part of speech tag, *noun* in the case of *au pair* and *hokus pokus*. The last word will also have regular morphological

features, while the preceding words will only have the the *ufl* feature. In the syntactic analysis, such phrases get a flat structure, see 4.5.

1.2.6.3 Idiomatic PPs

Quite a few prepositional phrases are idiomatic, such as *i fjor*, 'last year' and *med rette*, 'deservedly so'. If the prepositional complement is a word that is in use in the language apart from such expressions, it will receive its usual morphological tags. This is the case for *rette* in *med rette*. If the complement is a word which only occurs in such idiomatic prepositional phrases, we add the part of speech *noun* and the features *appell* (*appellative*) and *ubøy* (*ubøyelig*, 'undeclinable'), a feature which OBT uses in other cases.

- (1.18) *i fjor*
 form: i *lemma:* i *pos:* prep
 form: fjor *lemma:* fjor *pos:* subst *features:* appell|ubøy

- (1.19) *med rette*
 form: med *lemma:* med *pos:* prep
 form: rette *lemma:* rett *pos:* adj *features:* be|ent|pos

In the syntax we do not use a flat structure in such cases, but rather the normal syntactic analysis of transitive prepositions, cf. 4.3.

1.2.7 *Selv*

OBT has two morphological analyzes for *selv*, 'self': It can either be a determiner or an adverb. There is often not obvious when to choose one or the other analysis. We have therefore set up the criteria below. These criteria are based on the choices which seem to be implied by OBT's analysis, and they do not follow the analysis of *selv* in Faarlund et al., 1997, 920-922.

For the syntactic analysis of *selv*, see 5.16.

1. If *selv* does not precede a nominal constituent, it is an adverb

Selv is therefore an adverb in (1.20 a) and (1.20 b):

- (1.20) (a) *Vanskelig tilgjengelig er han ikke, selv om hans*
 hard accessible is he not self if his
 setningskonstruksjoner, rike på bisetninger, er lange.
 sentence constructions rich on subordinate clauses are long
 'He is not inaccessible, even though his sentences, rich on
 subordinate clauses, are long. (Aftenposten 012)
- (b) *Jeg føler meg også selv litt smittet.*
 I feel myself also self somewhat infected
 'I also feel a bit infected, myself.' (Aftenposten 012)

This principle also applies when *selv* follows a nominal constituent with which it does not semantically connected, as in (1.21):

- (1.21) *Norge som «humanitær stormakt» forskriver gjerne etiske*
 Norway as humanitarian superpower prescribes gladly ethical
medisiner for andre, men vil ikke ta den selv.
 drugs for others but will not take it self
 'As a "humanitarian superpower", Norway prescribes ethical drugs to
 others, but does not want to take them herself.' (Aftenposten 011)

2. If *selv* is placed after a nominal expression with which it is semantically connected, it should be tagged as a determiner In (1.22 a), *selv* is placed after *han*, and *han* and *selv* belong together semantically in a different way than in (1.21). *Selv* shall therefore be tagged as determiner. The same applies to (1.22 b) and (1.22 c)².

- (1.22) (a) *Ikke bare har det vært et problem, som han selv*
 not only has it been a problem which he self
skriver i forordet, å bli klok på Borten.
 writes in preface+the to become wise on Borten
 'Not only has it been a problem, as he himself writes in the preface,
 to fully understand Borten' (Aftenposten 013)
- (b) *Når nøden banker på, hjelper alle seg*
 when distress+the knocks on helps everyone refl.pron.
selv først.
 self first
 'When distress knocks on the door, everyone helps himself first.'
 (Aftenposten 009)
- (c) *[...] Vi har intet å frykte bortsett fra frykten selv,*
 we have nothing to fear except from fear+the self
og vi har ingen tid å miste [...].
 and we have no time to loose
 'We have nothing to fear except fear itself, and we have no time to
 spare' (Aftenposten 009)

3. *Selv* with the meaning 'even' shall be an adverb, even when it precedes a nominal expression This applies to sentences such as (1.23):

- (1.23) *Å begrense utgiftsveksten i årene fremover vil*
 to limit spending growth+the in years+the forward will
innebære politiske prøvelser som selv en flertallsregjering
 imply political ordeals which self a majority government
vil ha store problemer med å takle.
 will have big problems with to handle
 'To limit the growth in spendings the coming years will imply political
 ordeals which even a majority government will have big difficulties
 handling.' (Aftenposten 007)

²Due to insufficient documentation in a previous edition of these guidelines, this has not been implemented consistently in all parts of the treebank

1.2.8 Floating quantifiers

The morphological and syntactic analysis of floating quantifiers, such as (1.24 a) and (1.24 b), is described in 5.17.

(1.24) (a)

*Vi kommer alle til å bli blåst ut som lys på
we come all to to become blown out as candles on
ei bursdagskake.*

a birthday cake

'We will all be blown out, as candles on a birthday cake' (Dagbladet 003)

(b)

Begge måtte dei bøte med livet.

both must they pay with life+the

'They must both pay with their lives' (Faarlund et al., 1997, 921)

1.3 Sentence division

We usually have a sentence division after punctuation such as full stop, question mark, exclamation mark, colon and the special character | (see 1.1.7). There are however some exceptions to this general rule:

If sentence-separating punctuation is followed by quotation marks or parentheses, the sentence division shall come after the quotation mark or parenthesis:

(1.25) (a) *Barnet skal hete Leah Isadora Behn.*"

child+the shall be called Leah Isadora Behn

'The child shall be called Leah Isadora Behn.'" (Aftenposten 004)

(b) *(Det siste spørsmålet var for øvrig ikke retorisk.)*

the last question+the was for remaining not rhetorical

'(The last question was, by the way, not rhetorical.)'(Aftenposten 009)

If a sentence-separating punctuation mark occurs in the middle of something that syntactically must be regarded as a sentence, we do not split. Often, but not always, the text that follows the delimiter, starts with a lowercase letter in such cases:

(1.26) (a) *Grovt sett kan vi si at kartdata kommer fra*

roughly seen can we say that map data come from

tre kilder: det offentlige, kommersielle selskaper og

three sources the public commercial companies and

fra privatpersoner.

from private persons

'Roughly speaking, we can say that map data comes from three sources: the public sector, commercial companies and private individuals.' (Blogg-bm 002)

- (b) *Etter et herlig kveldsmåltid, bestående av aspargessuppe*
 after a wonderful evening meal consisting of asparagus soup
og hjort (barna fikk hvert sitt Nestlèglass
 and deer children+the got each his/her own Nestlè glass
da, du kan jo ikke servere barn asparges!!
 then you can certainly not serve children asparagus
Hello, lissom!), sprayet vi barna med Antibac og
 hello sort of sprayed we children+the with Antibac and
sendte de til sengs, mens jeg og min kjære ble
 sent they to bed while I and my dear became
sittende i loungen og stirre på hverandre med vellyst
 sitting in lounge+the and stare at each other with lust
i blikket.
 in gaze+the

'After a wonderful evening meal, consisting of asparagus soup and deer (although the children got a Nestlè glass each, you cannot serve asparagus to children!! Hello!), we sprayed the children with Antibac and sent them to bed, while I and my beloved remained sitting in the lounge, staring at each other with lust in the eyes.' (Blogg-bm 001)

We would like to point out that there are some instances in the corpus where these exceptions have not been implemented. We have not been able to correct this afterwards, as the tools we use do not allow modification of sentence boundaries.

Part II

Syntax

Chapter 2

Introduction

2.1 Dependency Grammar

We have chosen Dependency Grammar as grammatical formalism for NDT. Dependency Grammar has several advantages from a language technology point of view (cf. Johannessen et al. (2011, 5-7)). It is also a relatively transparent and simple formalism, and it should therefore be possible to annotate relatively quick.

Dependency Grammar has two important characteristics:

1. Syntactic structure is represented in the form of asymmetrical relations.
2. These relations are between individual words, not syntactic phrases.

With asymmetrical relation is meant a relation between two words, X and Y, where the existence of X depends on the existence of Y, but not vice versa (cf. Haug 2010, 5). Such an asymmetric relation is called a *dependency* or a *dependency relation*. The dependent word is called a *dependent*, while the word the dependent is related to, is called the *head*. Let us illustrate this with the following example:

(2.1) *Per kjøper røde biler nå.*
Per buys red cars now
'Per buys red cars now'

Sentence (2.1) consists of the verb *kjøper* and a subject and an object. There is a dependency relation between *kjøper* and the subject *Per*: Being the subject of *kjøper* is the reason for existence of *Per* in the sentence.

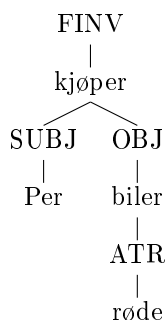
The object in (2.1) is a complex noun phrase, *røde biler*. Dependency relations are, however, between single words, not syntactic phrases. *Biler* is therefore directly dependent on the verb, while *biler*, in turn, serves as head of *røde*.

The verb *kjøper* is not itself dependent on any other word, and can therefore be said to be the head of the whole sentence. The finite verb of the matrix clause will always be the head of the whole sentence, according to our analysis. Often sentences without a finite verb occur, however. Then another word serves as to head. This will be further described in 3.1.

Dependency structure can be represented graphically as a tree structure, where heads are located over dependents. The sentence's head is at the top of

the tree. Each dependent is marked with a label indicating the type dependency relation we are dealing with: SUBJ, DOBJ, ADV, ATR etc. Such a label is called a *function*. The head of the sentence will also have a function called the *root function*. For finite verbs, the root function will be FINV. The dependency analysis for (2.1) looks like this:

(2.2)



Not all syntactic relations are so obviously asymmetrical. Two cases are particularly problematic: the relationship between function words and lexical words and coordination (cf. Nivre 2005):

(2.3) (a) *Per har kjøpt biler.*

Per has bought cars

'Per has bought cars.'

(b)

Per og Kari kjøper biler.

Per and Kari buy cars

'Per and Kari buy cars.'

In (2.3 a) the lexical verb *kjøpt* co-occurs with the auxiliary verb *har*. Both contain grammatical information relevant to each other and it is difficult to see that the one is in a superordinate relationship to the other: The auxiliary carries information on tense and aspect, while the lexical verb specifies the semantic content and argument structure. In the case of coordination, as in (2.3 b), each of the conjuncts can in principle fill the grammatical function in the sentence: *Per* and *Kari* could each be the subject of the sentence. It is therefore difficult to argue that there is an asymmetric relationship between them. In Tesnière (1965), the founding text of Dependency Grammar, dependencies are not used for function words and coordination, but separate, symmetrical relations. In line with what is common in dependency treebanks, however, we have chosen to use dependency relations in all cases. The way we analyze the coordination is described in section 4.4. The relationship between lexical words and function words will be treated already in the next section, 2.2.4.

2.2 Fundamental principles of analysis

In this section we will briefly to explain a few fundamental principles of analysis in NDT.

2.2.1 No empty nodes

To represent certain types of ellipsis, some treebanks use empty nodes, i.e. nodes in the dependency tree that are not associated with any of the words in the sentence (see e.g. Haug 2010). This is, however, disadvantageous for a number of language technology purposes. Since the language technology community is the primary target group for our corpus, we have chosen not to use empty nodes. Each node will, in other words, be associated with a word in the sentence. In 5.1, we describe in more detail how we analyze elliptical constructions.

2.2.2 Only unique dependencies

In this corpus, each node has one, and only one, dependency. We do not use secondary dependencies, e.g. to mark control and binding relations, as they do in the PROIEL corpus (cf. Haug (2010, 48-53)).

2.2.3 Crossing dependencies

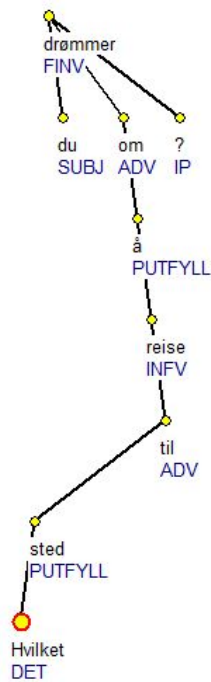
The dependency structure and the linear structure do not always fully match. A dependent may, in some constructions, be separated from the other dependents on the same head with one or more words. This occurs for example in wh-questions and topicalization. In wh-questions, the word order is often different from declarative clauses, as a wh-word or a phrase which contains a wh-word is “moved” to the left periphery of the clause. (cf. Faarlund et al. 1997, 936-943). In (2.4) it is the prepositional complement which is located in the left periphery.

- (2.4) *Hvilket sted drømmer du om å reise til?*
 what place dream you about to travel to
 ‘What place do you dream of traveling to?’ (Dagbladet 003)

If we shall provide an adequate analysis of such sentences, we will end up with syntactic constituents which do not constitute a single unit in the linear structure: There are several words between the prepositional complement *Hvilke ord* and its head, *til*. This phenomenon is sometimes referred to as *crossing dependencies*¹. Crossing dependencies may be problematic from a language technology point of view (cf. Nivre 2007). For us it is important that the analyzes are as linguistically adequate as possible, and we therefore allows crossing dependencies. The moved constituent in (2.4) should therefore be dependent on the preposition *til*, as the analysis in (2.5) shows:

- (2.5)

¹The term is slightly unintuitive with the visualization of dependency trees which we use here. In a dependency visualization which takes into account the linear order to a larger degree, it will be clearer that dependents intersect.



Crossing dependencies also occur in topicalization. In (2.6 a) is the subject of a clause topicalized and “moved” to the left periphery of the matrix clause. In our dependency analysis, the topicalized word is a dependent on the subordinate clause’s verb, as (2.6 b) shows:

- (2.6) (a) *Utsmyningsordningen* [...] *mener Hauge er problematisk*
 embellishment scheme+the thinks Hauge is problematic
av samme grunn.
 of same reason
 ‘The embellishment scheme Hauge thinks is problematic for the same reason.’ (Klassekampen 001)
- (b)



2.3 Overview of the syntactic tagset

FINV: Finite verb. See 3.1.1.

- FLAT:** Flat structure. See 4.5.
- FOBJ:** Formal object. See 3.2.4.
- FOPRED:** Free object predicative. See 3.8.
- FRAG:** Fragment. See 3.1.3.
- FSPRED:** Free subject predicative. See 3.8.
- FSUBJ:** Formal subject. See 3.2.2.
- IK:** Sentence-internal punctuation. See 5.2 and 5.11.
- INFV:** Non-finite verb. See 3.7.
- INTERJ:** Interjection. See 3.1.2 and 3.9.
- IOBJ:** Indirect object. See 3.4.2.
- IP:** Sentence-separating punctuation. See 5.2.
- KONJ:** Conjunction. See 4.4.
- KOORD:** Coordination. See 4.4.
- KOORD-ELL:** Coordination with verbal ellipsis. See 5.1.3.
- OPRED:** Object predicative. See 3.8.
- PAR:** Parenthetical expression. See 5.3.
- POBJ:** Potential object. See 3.2.4.
- PSUBJ:** Potential subject. See 3.2.3.
- PUTFYLL:** Prepositional complement. See 4.3.1.
- SBU:** Complementizer. See 3.10.2.
- SPRED:** Subject predicative. See 3.8.
- SUBJ:** Subject. See 3.2.
- UKJENT:** Superfluous words. See 5.4.

Chapter 3

Sentences

3.1 Root functions

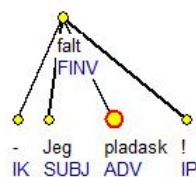
Root functions are the functions used for the highest head in a dependency structure. We use three root functions: FINV (finite verb), INTERJ(ection) and FRAG(ment). These functions are presented respectively in 3.1.1, 3.1.2 and 3.1.3. 3.1.4 describes how we analyze sentences with auxiliary verbs.

3.1.1 FINV

If a dependency structure is a finite sentence, as in (3.1 a), the finite verb of the matrix clause should always be the head. The root function for finite verbs is FINV, cf. (3.1 b):

- (3.1) (a) - *Jeg falt pladask!*
I fell plump
'I fell in love at first sight!' (Dagbladet 003)

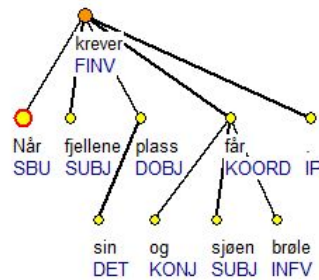
(b)



Sometimes seemingly subordinate clauses appear as independent sentences. (3.2 a) exemplifies this. In such cases, we also use FINV, cf. (3.2 b):

- (3.2) (a) *Når fjellene krever sin plass og sjøen*
when mountains+the demand their space and sea+the
får brøle.
is allowed to roar
'When the mountains demands space and the sea is allowed to roar.'
(Dagbladet 003)

(b)



It is important to emphasize that FINV should only be used as a root function. Finite verbs in subordinate clauses shall always carry the clause's grammatical function. This is described in section 3.10. In the case of coordinated matrix clauses, the function KOORD is used, cf. 4.4.

3.1.2 INTERJ

Interjection can often stand alone, without any finite verb, as in (3.3).

(3.3) - *Ja.*
yes

'Yes.' (Dagbladet 003)

While dependency structures with nouns, prepositions, etc. as head often are fragments that imply a verb of some kind, an interjection can constitute a complete meaning in itself. Therefore, we have chosen to let INTERJ be a root function¹. The analysis of (3.3) is thus as follows:

(3.4)



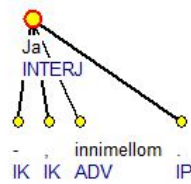
Interjections in root position can take dependents (in addition to punctuation). In most cases it will be ADV dependents, see (3.5 a) and the analysis in (3.5 b):

(3.5) (a) - *Ja, innimellom.*
yes sometimes

'Yes, sometimes.' (Dagbladet 003)

(b)

¹INTERJ may also be used on dependents on verbs, cf. 3.9.



If an interjection occurs in an independent sentence, the finite verb, not the interjection, should be head. For more on this, see 3.9.

3.1.3 FRAG

If it is possible to choose the root functions FINV or INTERJ in a sentence, we do that. When this is not possible, we use the root function FRAG. (3.6 a) and (3.6 b) are examples of such sentences:

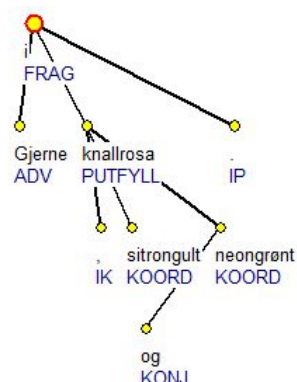
- (3.6) (a) *Klar beskjed.*
 clear message
 'Clear message.' (Dagbladet 001)
- (b) *Gjerne i knallrosa, sitrongult og neongrønt.*
 often in shock pink lemon yellow and neon green
 'Often in shock pink, lemon yellow and neon green.' (Dagbladet 002)

Since FRAG is used in very different cases, there is no generally applicable rule for deciding which word should receive the root function. It must be decided in each case what word is most natural to choose as head. In (3.6 a) *beskjed* must be head, because it is not itself dependent on any other word, while *klar* modifies *beskjed*, see (3.7 a). There are good reasons to think that *gjerne* modifies the preposition in (3.6 b). The preposition itself is not dependent on any other word. It therefore becomes the head, cf. (3.7 b).

- (3.7) (a)



- (b)



Usually we use the same functions for dependents to words with the FRAG function as we would use if the same words had appeared as part of a complete sentence. But sometimes it is necessary to imply a verb to get a head for some of the words in the fragmentary sentence. This applies for example in (3.8). The prepositional phrase *uten at passet er med* is probably dependent on an implied verb, eg. *går*, 'goes', not on the noun *dag*.

- (3.8) *Aldri en dag uten at passet er med.*
 never a day without that passport+the is with
 'Never a day without bringing the passport' (Dagbladet 001)

In 5.1 we explain how such sentences should be analyzed.

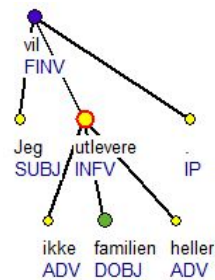
3.1.4 The relationship between auxiliary verbs and lexical verbs

A consequence of the fact that the finite verb in a matrix clause always bears the root function, is that auxiliary, not the lexical verb must be the head, since it is always the auxiliary which is finite (for the relationship between functional and lexical words in general, see 2.2). The lexical verb is a dependent on the auxiliary verb with the function INFV (non-finite verb; see also 3.7). In (3.9), it is the auxiliary *vil* that will be the head, while the lexical verb *utlevere* gets the function INFV:

- (3.9) *Jeg vil ikke utlevere familien heller.*
 I will not expose family+the after all
 'After all, I don't want to expose my family.' (Dagbladet 003)

In constructions with auxiliary verbs, there are in principle two possible verbs on which you can attach dependents. We have decided that subjects should be dependents on the finite verb, since tense usually is a necessary condition for having overt subjects. All other arguments and verbal adjuncts will be dependent on the lexical verb. In (3.9) the subject, *Jeg*, will be dependent on the auxiliary *vil*, while the direct object *familien* and the two adverbials *ikke* and *heller* are dependents on the lexical verb *utlevere*, as the analysis in (3.10) shows:

(3.10)



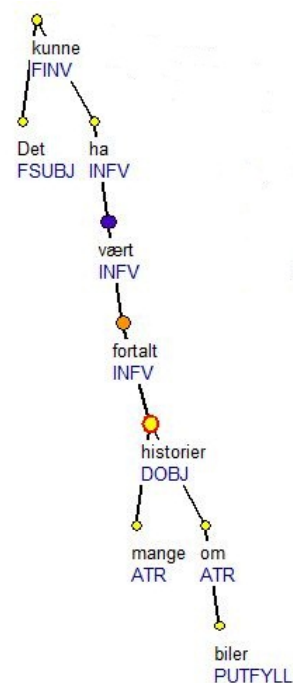
Conjunctions, complementizers and sentence-separating punctuation are dependents on the finite verb. Sentence-internal punctuation will often be dependent on the finite verb, see 5.2. Formal subjects will be dependents on the finite verb, while potential subjects will be dependents on the lexical verb (cf. 3.2).

Sometimes chains of auxiliary verbs occur, as in (3.11):

- (3.11) *Det kunne ha vært fortalt mange historier om biler [...].*
 it could have been told many stories about cars
 'Many stories could have been told about cars.' (Dagbladet 003)

In such cases, the finite auxiliary is the head. Each of the following auxiliaries are dependents on the preceding ones. The lexical verb is dependent on the last auxiliary. The analysis of (3.11) is thus as follows:

(3.12)



In subordinate clauses, the relationship between auxiliaries and lexical verbs is the same as what we have outlined here. In that case, however, the finite verb has the grammatical function of the subordinate clause, not FINV. For more on subordinate clauses, see 3.10.

3.2 Subjects

This subsection explains the different types of subjects and subject functions used in NDT: referential subjects in subject position (SUBJ), formal subjects (FSUBJ) and potential subjects (PSUBJ). We also include a brief description of formal and potential objects (FOBJ and POBJ), since these functions can best be understood in connection with their subject counterparts.

3.2.1 Referential subjects in subject position (SUBJ)

We will not try to give an exhaustive description of subjects in Norwegian here, but content ourselves with a brief characterization (see for example Faarlund et al. (1997, 674-693) for a more detailed presentation). The subject can be defined using word order criteria: In sentences with matrix clause word order, subjects are either in the position right before or right after the finite verb, as in (3.13a) and 3.13b). In sentences with subordinate clause word order, the subject is before the finite verb, but not necessarily immediately before - the finite verb and the subject can be separated by a sentence adverbial, as in (3.13c).

- (3.13) (a) *Gjestene forsynte seg godt av buffeten.*
 guests+the provide refl good from buffet+the
 'The guest took their fill at the buffet.'
- (b) *Etter maten tok de seg en lur.*
 after food+the took they refl a nap
 'After the meal, they took a nap.'
- (c) *Gi beskjed hvis du ikke kan komme likevel.*
 give message if you not can come after all
 'Tell me if you cannot come after all.'

The positions mentioned here, we refer to as (possible) subject positions.

In our analysis we will distinguish between referential and non-referential subjects, and referential subjects in subject position will have the function SUBJ. Non-referential subjects (i.e. semantically empty formal subjects), are also in subject position, but will have a different function. Formal subjects are discussed below in 3.2.2.

Referential subjects in subject position often have the form of nominal expressions, but they can also be clauses or infinitives, as we see in the examples below:²

²Sometimes, probably in more marginal cases, prepositions may also be subjects. This applies to *her*, 'here' and *der*, 'there' (see 3.2.2 with note 3), but it may also apply in other cases (see Faarlund et al. (1997, 682-683)).

- (3.14) (a) *Allerede i går kom de første meldingene om*
 already yesterday came the first reports+the about
kaos på veiene.
 chaos on roads+the
 'Already yesterday, the first reports about chaos on the roads came.'
- (b) *At du har gjennomført dette prosjektet, er en stor*
 that you have completed this project+the is a big
bragd.
 achievement
 'That you have completed this project, is a big achievement.'
- (c) *Å trene er bra for både kropp og sjel.*
 to exercise is good for both body and soul
 'To exercise is good for both the body and the soul.'

In sentences with both finite and infinite verbs, words with a SUBJ relation (and also FSUBJ, but not PSUBJ) should be dependent on the finite verb, unlike other arguments and adjuncts (see 3.1.4). We have chosen this analysis because finiteness and overt subjects correlate in Norwegian.

3.2.2 Formal subjects (FSUBJ)

Formal subjects do not have any clear reference, unlike the subjects described above in 3.2.1. Formal subjects are elements with no semantic content, that are tasked to fulfill the general requirement in Norwegian that finite sentences must have an overt subject. The formal subject's form is most commonly *det*, but also *der* may occur in some constructions (mainly in presentational sentences).³ We give formal subjects the function FSUBJ. Here follows a rough overview of structures with formal subjects.

3.2.2.1 Weather sentences and related constructions

Formal subjects occur in weather sentences and some related constructions (i.e. constructions where the predicate distributes a so called quasi-argument, see e.g. Falk (1993, 67-90)). The following examples illustrate this.

- (3.15) (a) *Det blåser fælt ute.*
 it blows horribly outside
 'The wind blows horribly outside'
- (b) *Hvordan går det med deg?*
 how goes it with you
 'How are you?'
- (c) *Det er dårlige tider.*
 it are bad times
 'Times are bad' (Faarlund et al., 1997, 679)

³In some dialects the form *her*, 'here', is used in sentences such as *Her er kaldt*, 'it is cold'. However, *here* always has a certain locative content (Faarlund et al., 1997, 681-682), and we will not count it as a formal subject. Also, *der* can in similar cases (where there is no other subject) have a clear locative meaning, 'there'. The most linguistically appropriate is likely to count *her* and *der* in such constructions as referential entities, i.e. SUBJ, even though they are not nominal.

3.2.2.2 Presentational sentences

We find formal subjects in presentational sentences. From an information structure point of view, presentational sentences are characterized by the fact that they introduce an NP with new information in the discourse. As new information is not preferred in subject position in Norwegian, this nominal phrase is in object position, while the subject position is filled by a formal subject.

- (3.16) (a) Det står en mann i hagen.
 it stands a man in garden+the
 'There is a man in the garden.'
- (b) Det er nok mat i kjøleskapet.
 it is enough food in fridge+the
 'There is enough food in the fridge.'
- (c) I går kom det en ny elev i klassen.
 yesterday came it a new pupil in class+the
 'Yesterday, a new pupil came in class.'

The nominal phrase with new information which is in object position, is called a potential subject and receive the function PSUBJ. We discuss this function in more detail in 3.2.3 below. Occasionally, presentational sentences are difficult to distinguish from other structures with formal subject, see 3.2.6 for more on this. Note also that presentational sentences is one of the constructions where *der* may act as formal subject instead of *det*.

3.2.2.3 Impersonal passive sentences

Passive means that a verbs subject is “deleted”. This allows for an object to take over the subject function, as below, where *mange bøker* has become the subject:

- (3.17) Mange bøker har blitt skrevet om dette fenomenet.
 many books have been written about this phenomenon+the
 'Many books have been written about this phenomenon.'

In impersonal passive sentences, however, the object remains, while a formal subject fills the subject position:

- (3.18) Det har blitt skrevet mange bøker om dette fenomenet.
 it has been written many books about this
 phenomenon+the
 'There have been written many books about this phenomenon.'

3.2.2.4 In clause-anticipating constructions:

Clause-anticipating constructions contain a clause or infinitive construction which semantically can be regarded as the subject, but which is not in the subject position. The subject position is filled by the formal subject. Clause-anticipating constructions often contain a predicative adjective, *fin* and *morsom* in the examples below:

- (3.19) (a) *Det var fint at du kunne komme.*
 it was nice that you could come
 'It was nice that you could come.'
- (b) *Etter en bedre middag er det ikke så morsomt å vaske*
 after a better dinner is it not so fun to wash
opp.
 up
 'After a good dinner, it isn't so much fun to wash up.'

We give the clause/infinitive in clause-anticipating constructions the function potential subject (PSUBJ), see 3.2.3.

3.2.2.5 Clefting

The examples below illustrate clefting:

- (3.20) (a) *Det er jeg som har gjort det.*
 it is I who have done it
 'It is me who have done it'
- (b) *Det er noen som misliker hunder.*
 it is someone who dislike dogs
 'There are some people who dislike dogs'

We operate with two slightly different types of clefting, both containing a formal subject. See 3.3 for the description of the analyzes.

3.2.3 Potential subjects (PSUBJ)

Potential subjects are referential, but do not stand in any of the usual subject positions, in contrast to the subjects discussed in 3.2.1. In structures with potential subject, a formal subject stands in one of the subject positions, while the potential subject is in a position where we usually find objects. Potential subjects receive in our annotation the function PSUBJ.

Potential subjects occur together with formal subjects in presentational sentences and clause-anticipating constructions. In presentational sentences, the potential subject is a nominal phrase:

- (3.21) (a) *Det står en mann i hagen.*
 it stands a man in garden+the
 'There is a man in the garden.'
- (b) *Det er nok mat i kjøleskapet.*
 it is enough food in fridge+the
 'There is enough food in the fridge.'

In clause-anticipating constructions, the potential subject is a clause or an infinitive. The clause-anticipating constructions often contain a predicative as well as the potential subject (respectively *fint*, 'good', and *så morsomt*, 'so funny', in the examples below):

- (3.22) *Det var fint at du kunne komme.*
 it was nice that you could come

- (3.23) *Etter en bedre middag er det ikke så morsomt å vaske*
 after a better dinner is it not so funny to wash
 opp.
 up
 'After a good dinner, it is not so much fun to wash up.'

Sometimes we find an object instead of a predicative, such as *ingenting*, 'nothing', in (3.24). We still count the clause *om du er sen* as a potential subject.

- (3.24) (3.25) *Det gjør ingenting om du blir sen.*
 it does nothing if you become late
 'It doesn't matter if you are late.'

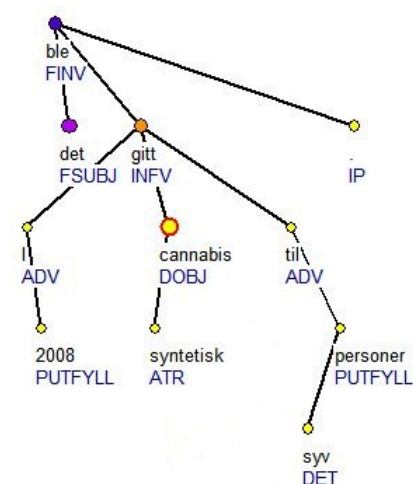
Potential subjects can often (but not always) be made the subject if the formal subject of the sentence is deleted. Example (3.26) below shows that this is a possible operation (3.22) above:

- (3.26) *At du kunne komme, var fint.*
 that you could come was nice
 'That you could come, was nice.'

Note that we do not use PSUBJ by impersonal passive sentences, as in (3.27 a). In such sentences, we assume that the referential argument is a direct object (cf. 3.2.2.3), and we therefore use the function DOBJ:

- (3.27) (a) *I 2008 ble det gitt syntetisk cannabis [...] til*
 in 2008 became it given synthetic cannabis to
syv personer.
 seven persons
 'In 2008, synthetic cannabis was given to seven people.'
 (Aftenposten 011)

(b)



Sometimes it can be difficult to mark potential subjects off from other functions. See 3.2.5 and 3.2.6 for a description of how we distinguish potential subjects (or objects) from adverbials and subject predicatives, respectively.

3.2.4 Formal and potential *objects* (FOBJ and POBJ)

As we indicated above, we operate with formal and potential objects in addition to the above-mentioned formal and potential subjects. Formal and potential objects have very much in common with formal and potential subjects and issues we face in connection with formal/potential subjects may be relevant to formal/potential objects (see for example 3.2.5 below). We will therefore briefly mention formal and potential objects here before we move on.

Formal objects have the function FOBJ. Formal objects take the form of a semantically empty *det*, as formal subjects do, but is in an object-relation to a verb, not a subject relation. The following example illustrates this (the formal object is emphasized):

⁴

- (3.28) *Det var Harald Zwart som på 1990-tallet fant det helt*
 it was Harald Zwart that on 90ies found it entirely
 nødvendig å mobbe norsk film.
 necessary to bully Norwegian film
 ‘It was Harald Zwart who, in the 90ies, found it entirely necessary to
 bully Norwegian film. (Dagbladet 001)’

Det has an object-relation to *fant*, and therefore cannot be considered a subject, but at the same time we see that the construction is parallel to the clause-anticipating constructions we mentioned above in 3.2.2.4. If we replace the verb *fant* with *være*, ‘be’, we can easily rewrite the relevant part of the sentence to a clause-anticipating construction with a formal subject:

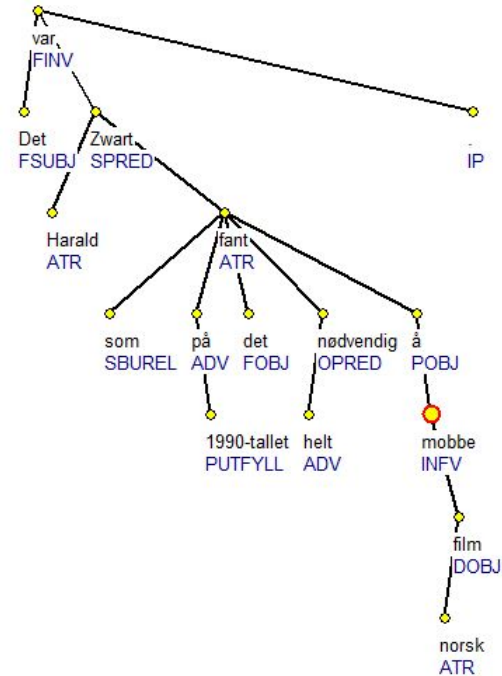
- (3.29) (a) *Det er helt nødvendig å mobbe norsk film.*
 it is entirely necessary to bully Norwegian film
 ‘It is entirely necessary to bully Norwegian film.’

The example in (3.28) illustrates not only formal objects, but also potential objects, which we will give the the function POBJ. The infinitive in the example sentence, *å mobbe norsk film*, is entirely parallel to the potential subject in a clause-anticipating construction with a formal subject, but the whole construction is in an object relative to *fant*, and the infinitive is therefore a potential object instead of a potential subject. Below, (3.28) is repeated (the formal object is emphasized), and in (3.31 a) we see the full dependency analysis of the sentence.

- (3.30) *Det var Harald Zwart som på 1990-tallet fant det helt*
 it was Harald Zwart that on 90ies found it entirely
 nødvendig å mobbe norsk film.
 necessary to bully Norwegian film
 ‘It was Harald Zwart who, in the 90ies, found it entirely necessary to
 bully Norwegian film. (Dagbladet 001)’

⁴In this example, the formal object could be analyzed as a direct object. Choices we have made concerning constructions with object and infinitive, imply that we could need a function for formal *indirect* objects in some marginal cases (3.4.3) for example in a sentence such as *Gud lot det regne*, ‘God made it rain’. However, we will not introduce a separate function with such a limited application, and we therefore let the function FOBJ cover all kinds of formal objects.

(3.31) (a)



3.2.5 How to distinguish between potential subjects/objects and adverbials

Sometimes, prepositional phrases serve as potential subjects/objects. This is particularly true for prepositional phrases headed by *med*, 'with':

(3.32) (a) *Det er hyggelig med besøk.*

it is nice with visit

'It is nice with a visit.'

(b) *Vi reknar det for altfor dyrt med drosje.*

we consider it for way too expensive with taxi

'We consider it way too expensive to take a taxi' (Faarlund et al., 1997, 1026)

(c) *Frå Kommunerevisjonen straumar det på med meldingar*

from Municipality Audit streams it on with messages

om kortsynt planlegging, manglande oversikt og

about short-sighted planning lacking perspective and

urovekkjande naivitet.

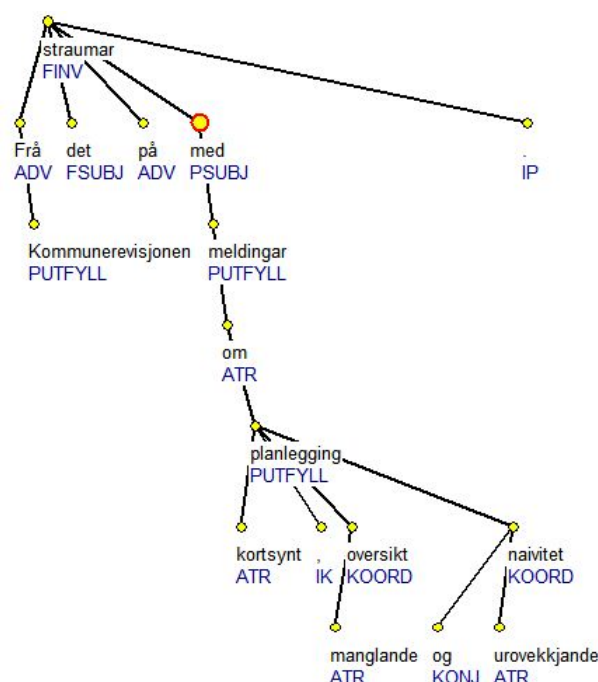
troubling naivety

'From the municipality audit it swarms with messages about

short-sighted planning, lacking perspective and troubling naivety.'

(Klassekampen-nn 003)

(d)



These can sometimes be difficult to distinguish from prepositional phrases which have an adverbial function. How should *med røntgenmaskinen* in the example below be analyzed?

- (3.33) *Ha det moro med røntgenmaskinen.*
 have.imp it fun with x-ray machine+the
 'Have fun with the x-ray machine.' (Dagbladet 001)

We use a substitution test to distinguish the constructions apart. If the prepositional phrase can be replaced by an infinitive or a clause with the complementizer *at*, we consider it a potential subject/object. If this is not possible, we consider it an ADV-dependent. In other words, we must have clear evidence before using the functions PSUBJ/POBJ. If we don't have such evidence, we use ADV. We see that the prepositional phrase in (3.32 a) above can be replaced with an infinitive/an *at*-clause, and we therefore choose PSUBJ:

- (3.34) *Det er hyggelig å få besøk / at folk kommer på besøk.*
 it is nice to get visit / that people come on visit
 'It is nice to get visitors/that people come visiting.'

This is not possible in (3.33)⁵:

- (3.35) **Ha det moro å leke med røntgenmaskinen.*
 have it fun to play with x-ray machine+the
 '*Have fun to play with the x-ray machine.'

⁵Ungrammatical sentences are marked with an asterisk

We therefore analyze (3.33) as ADV. The same applies in sentences such as (3.36 a) below:

- (3.36) (a) *Det vrimlar med folk i gatene.*
 it swarms with people in streets+the
 'It swarms with people in the streets.' (Faarlund et al., 1997, 837)
- (b) **Det vrimlar at det er folk i gatene.*
 it swarms that it is people in streets+the
 '*It swarms there are people in the streets.'

(In the last example, our analysis differs from NRG, which considers *med folk* in (3.36 a) as a potential subject. These guidelines are not the ideal place for a discussion of this issue, but we believe that predicates like *vrímle*, 'swarm', are of a different type than predicates that can take a formal subject and infinitive/*at*-clause as potential subject.)

3.2.6 How to distinguish between potential subjects and subject predicatives

Sometimes it can be difficult to distinguish sentences with formal and potential subject from sentences with formal subject and subject predicative, especially with the verb *være*, but also in some other case. How should we analyze the emphasized phrases in the examples below?

- (3.37) (a) *Det er et stort juletre på torget.*
 it is a big Christmas tree on square+the
 'There is a big Christmas tree in the square.'
- (b) *Det er bare meg!*
 it is just me
 'It is just me'
- (c) *23. september 1995 ble det bryllup på en*
 23d September 1995 became it wedding on a
 fjelltopp i Gudbrandsdalen
 mountain peak in Gudbrandsdalen
 'September 23d 1995, there was a wedding on a mountain peak in
 Gudbrandsdalen.' (Dagbladet 001)
- (d) *Det er bare å håpe på bedring.*
 it is just to hope on improvement
 'We can only hope for improvement'

It is difficult to find clear-cut rules that solve the problem of separating the two functions apart, but we will rely on the criteria below in our analysis. As we saw in 3.2.3, we only use the function PSUBJ in presentational sentences and clause-anticipating constructions - in presentational sentences, the potential subject is a nominal constituent, in clause-anticipating constructions it is a clause or an infinitive.

Not all the following criteria apply to all cases of doubt - the criteria in 1- 3 are only relevant for (possible) presentational sentences, not (possible) clause-anticipating constructions. Some criteria may also give vague or contradictory

results. We therefore choose predicative in cases where the other criteria do not provide good answers (see 5).

The criteria are as follows:

1. *Properties of the verb in presentational sentences.*

- We do not believe that presentational sentences can have *bli* as their verb. If the verb is *bli*, we use predicative (see Faarlund et al., 1997, 734). Therefore, *bryllup...* in (3.37 c) should be analyzed as predicative.
- If a presentational sentence has *være* as its verb, will it be possible to substitute *være* with *finnes*, 'be found', or *eksistere*, 'exist' without the meaning changing significantly. We see that this is possible in (3.37 a) above, which we will analyze as a presentational sentence with a potential subject:

(3.38) *Det finnes et stort juletre på torget.*
 it is found a big Christmas tree on square+the
 'A big Christmas tree is found in the square.'

If *være* is not synonymous with *finnes* or *eksistere*, but rather is a copular verb (i.e. almost meaningless), we analyze it as a predicative (see Faarlund et al., 1997, 734). (3.37 b) above is a striking example of this:

(3.39) **Det fins / eksisterer bare meg!*
 it is found / exists only me
 'It is only me!'

2. *Obligatory adverbial in presenting sentences in which være is synonymous with finnes.* If the verb in a presentational sentence is *være*, and *være* can be replaced with *finnes*, an adverbial will normally be required (see Faarlund et al., 1997, 830)⁶. In example (3.37 a) the place adverbial probably cannot be omitted (assuming that *det* still is a formal subject and does not refer to anything outside of the sentence):

(3.40) *??Det er et stort juletre.*
 it is a big Christmas tree
 '*There is a big Christmas tree.'

3. *Indefiniteness.* Potential subjects in Norwegian must generally be indefinite, and it will therefore often be appropriate to consider definite nominal expressions, pronouns, proper nouns and phrases with possessives and determinatives as predicatives⁷. This criterion is particularly relevant for example (3.37 b), repeated below:

(3.41) *Det er bare meg!*
 it is just me
 'It is just me'

⁶This does not apply if *være* is synonymous with *eksistere*, then it can stand alone.

⁷There are, however, some exceptions to this rule, in particular in constructions with *bare*, 'just', and superlative.

As *meg* is a pronoun, we consider it a predicative, not a potential subject.

4. *Movement test.* If it is possible to drop the formal subject and make the appropriate nominal phrase/subordinate clause/infinitive the subject, it is a potential subject (see 3.2.3 above). (Faarlund et al., 1997, 835).⁸ This test sometimes gives somewhat inconclusive results for nominal phrases, as in (3.37 a):

(3.42) *Et stort juletre er på torget.*
 a big Christmas tree is on square+the
 'A big Christmas tree is in the square.'

But this probably has an independent explanation in the fact that subjects in Norwegian preferably should be specific in one sense or another, and we will therefore stick to the analysis of *et stort juletre* as a potential subject, as the tests above indicate. For clause-anticipating constructions, the movement test often gives a clearer result. (3.37 d) above, repeated here, may at first glance resemble a clause-anticipating construction, but note that it does not contain any predicative next to the infinitive, as clause-anticipating constructions usually do (we analyze *bare* as an adverbial):

(3.43) *Det er bare å håpe på bedring.*
 it is just to hope on improvement
 'We can only hope for improvement'

We also see the movement test does not work:

(3.44) **Å håpe på bedring er bare.*
 to hope on improvement is just
 '*To hope for improvement is just.'

We can see this as an indication that the sentence should be analyzed with predicative, although the movement test should be used with reservation.

5. *In case of persistent doubt: Select predicative.* If the rules above do not help, select predicative. Predicative is a function that occurs in different constructions, while potential subject, as defined by us, has a more limited distribution. It seems reasonable to require a clearer justification for potential subjects than for predicative. Of the examples we brought up in the introduction to this subsection, it is the example (3.37 d) which is the most difficult to classify using the other criteria, although we saw the movement test in 4 suggested predicative. Since there are no strong arguments against this, we classify *å håpe på bedring* as predicative, based on this rule.

3.3 Clefting

So-called *clefting* is a common type of construction in Norwegian. We have four examples below:

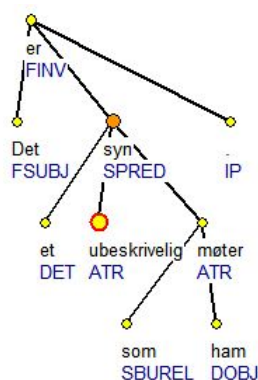
⁸There are potential subjects that, for various reasons, do not pass this test. But it is probably correct to say that all phrases that do pass it, are potential subjects.

- (3.45) (a) *Det er et ubeskrivelig syn som møter ham.*
 it is an indescribable sight which meets him
 'It is an indescribable sight which meets him.' (Dagbladet 001)
- (b) *Det var i går den store begivenheten skjedde.*
 it was yesterday the big event+the happened
 'It was yesterday that the big event took place.'
- (c) [...] *det er ingen som viser mer NHL i Europa enn oss.*
 it is none who shows more NHL in Europe than us
 'There are none who show more NHL in Europe than us.'
 (Dagbladet 001)
- (d) *Det er noen som misliker hunder.*
 it is some who dislike dogs
 'There are some who dislike dogs.'

Phrases with clefting contain a formal subject (see 3.2.2 for more on this), the verb *være* and a relativized phrase (*ubeskrivelig syn* and *ingen* in the examples above) moved up from a subordinate clause which resembles a relative clause. The clause does not behave quite like other relative clauses, but we still decided to analyze it with ATR/ADV in the same way as more prototypical relative clauses (see 3.10.3 for more on this).⁹

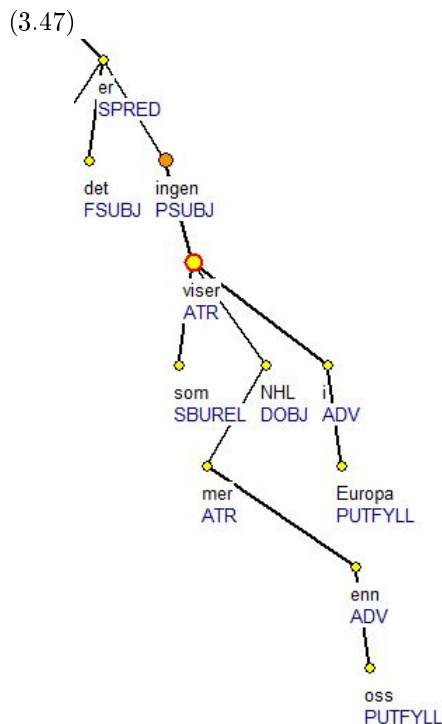
Clefting comes in two versions, the focus type and the presentational type (Faarlund et al., 1997, 1088). The examples (3.45 a) and (3.45 b) above are of the focus type, while (3.45 c) and (3.45 d) are of the presentational type. As is evident from the names, the focus type is used to focus information, while the presentational type can be seen as a type of presentational sentence (see 3.2.2 for more on these). In line with Faarlund et al. (1997, 1088-1095), we will give a somewhat different analysis to the two constructions - the difference is the function the relativized phrase should have. In the focus type, we will give this phrase the function subject predicative (SPRED), in the presentational type it should however be potential subject (PSUBJ). Example (3.45 a), which is of the focus type, therefore gets this analysis:

(3.46)



⁹A characteristic of the "relative clause" in clefting is that it does not topicalize along with its correlate: * Et ubeskrivelig syn som møter ham, er det.

Example (3.45 c), of the presentational type, is analyzed as follows, with *ingen* as PSUBJ:



(The reason why the highest verb, *være*, has received the function SPRED, is of course that the example originally is part of a larger sentence.)

Sometimes it can be difficult to distinguish the two types. We use the following criteria to determine which analysis we choose (based on Faarlund et al. (1997, 1088-1095) and Lundeby (1976)):

1. *Stress*. When focused, the predicative becomes highlighted and focused, and therefore receives stress. The potential subject of the presentational type can, but need not necessarily receive stress. The examples (3.45 a) and (3.45 b) are difficult to imagine without stress on *ubeskrivelig syn* and *i går*, while *ingen* and *noen* in (3.45 c) and (3.45 d) can easily be unstressed.
2. *Presupposition*. In the focus type the content of the “relative sentence” is presupposed. The subordinate clause’s content is, in other words, an implicit condition for the whole construction, or it is known from the preceding context¹⁰. In (3.45 a), it is, for example presupposed that someone was met by something. In (3.45 d) it is, on the other hand, not presupposed that someone dislikes something.
3. *Properties of the verb*. In the presentational type, *være* may well be replaced by *finnes/eksistere* without the sentence’s semantics changing significantly. We see that this is possible in (3.45 c) above, but not in (3.45 a):

¹⁰Texts can also start with a clefting construction of the focus type, but this can be understood as a stylistic effect (Lundeby, 1976).

- (3.48) (a) [...] *det finnes ingen som viser mer NHL i*
 it is found none who shows more NHL in
 Europa enn oss.
 Europe than us
 'There exist none who show more NHL in Europe than us.'
- (b) **Det finnes et ubeskrivelig syn som møter ham.*
 it is found an indescribable sight which meets him
 '*It is found an indescribable sight which meets him.'

4. *The formal subject's form.* Dialects that may have *der* as formal subject in presentational sentences, can have *der* also with clefting construction of the presentational type, but not in the focus type:

- (3.49) (a) *Der er ingen som viser mer NHL i Europa*
 there is none who shows more NHL in Europe
 enn oss.
 than us
 'There are none who show more NHL in Europe than us.'
- (b) **Der var i går den store begivenheten*
 there was yesterday the big event+the
 skjedde.
 happened
 '*There was yesterday that the big event took place.'

(This test presupposes of course some knowledge of one of the relevant dialects.)

5. *If in doubt - select predicative.* In the discussion of the distinction between subject predicative and potential subject (see 3.2.6), we set up a general which states that we choose predicative when in doubt. We will let this rule apply also in clefting constructions. Structures with clefting should therefore be analyzed as the focus type (i.e. with predicative) if the above criteria does not give any satisfactory answer to which subtype the sentence belongs.

It is important to remember that the relativized clause in focus-type clefting can be virtually any type of phrase, not only nominal arguments. In example (3.45 b), repeated here, is *i går* a prepositional phrase with an adverbial function in the subordinate clause:

- (3.50) *Det var i går den store begivenheten skjedde.*
 it was yesterday the big event+the happened
 'It was yesterday that the big event took place.'

Here, it may be tempting to use the function ADV on *i går*, but the correct analysis is SPRED.

3.4 Objects

In this section we will give an account of the object functions in NDT and how they are used. In 3.4.1, we discuss the function DOBJ (direct objects), in 3.4.2, the function IOBJ (indirect objects), and in 3.4.3, constructions with perception verbs and the verbs *be*, 'pray, ask' and *la*, 'let', with an object and an infinitive. In addition to the object functions DOBJ and IOBJ, We also operate with formal objects (FOBJ) and potential objects (POBJ). These may best be understood in the context of formal and potential subjects, and we have therefore chosen to place the discussion of them in section 3.2. See Faarlund et al. (1997, 693-731) for a more general account of objects in Norwegian.

3.4.1 Direct objects (DOBJ)

3.4.1.1 Complements to verbs

We use the DOBJ function for direct objects which are complements verbs. DOBJs can take the form of both nominal constituents, adjectives infinitives and clauses, but they can in most cases not be prepositions (in elliptical constructions, see 5.1). Some verbs take prepositions as arguments, but in our analysis we do not distinguish between these and adjunct prepositions (see 3.5). So-called prepositional objects will therefore not get the function DOBJ, but ADV.

We will analyze the emphasized phrases in the examples below as DOBJ:

- (3.51) (a) *Alvheim ledet sosialkomiteen også i forrige periode*
 Alvheim led social committee+the also in last period
 [...]

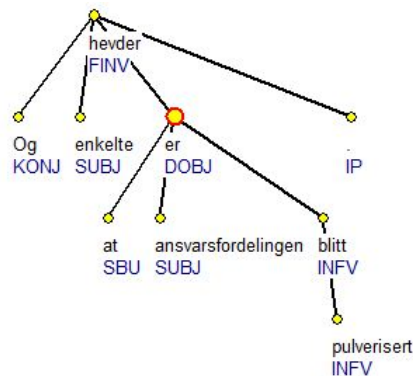
'Alvheim led the social committee also last period.' (Aftenposten 001)

- (b) *Litt forsinket, klokken 20.25, satte bryllupsgjestene seg til bords for å nyte gallamiddagen på Slottet.*
 a little late, clock+the 20.25 sat wedding guests+the
 refl to table for to enjoy gala dinner+the on palace+the
 'A little late, at 20:25, the wedding guests took a seat to enjoy the gala dinner at the palace.' (Aftenposten 001)

- (c) *Og enkelte hevder at ansvarsfordelingen er blitt pulverisert.*
 and some claim that division of responsibility+the is
 become pulverized
 " (Klassekampen 003)

In (3.51 a) we have a noun as DOBJ. In (3.51 b) we have the reflexive pronoun *seg* as DOBJ to *satte*, and also *gallamiddagen på slottet* as DOBJ to the infinitive *nyte*. (3.51 c) shows that subordinate clauses may also have the function DOBJ. The figure below shows the structure of 3.51 c:

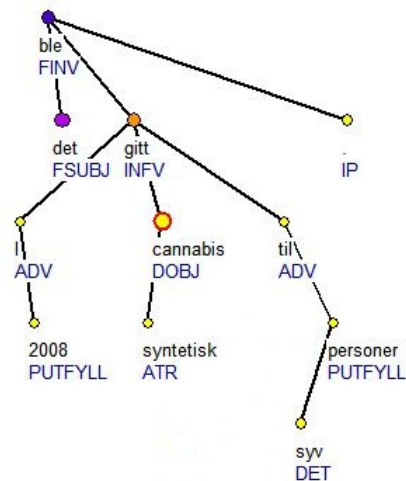
(3.52)



Note that we use DOBJ also in impersonal passive sentences, cf. 3.2.2.3 and 3.2.3:

- (3.53) (a) *I 2008 ble det gitt syntetisk cannabis [...] til syv personer.*
 in 2008 became it given synthetic cannabis to
 seven persons
 'In 2008, Synthetic cannabis was given to seven individuals.'
 (Aftenposten 011)

(b)

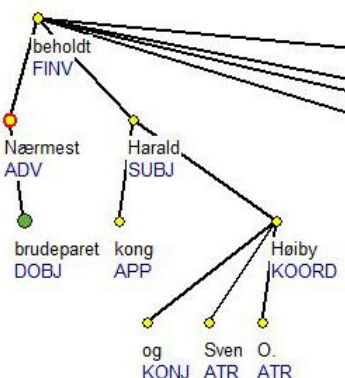


3.4.1.2 Complements to adjectives

A few adjectives can take a nominal complements in Norwegian - sometimes we encounter examples like *være redd hunder*, 'be scared [of] dogs', *bli kvitt pengene*, 'get rid [of] the money' *være nær noen*, 'be close [to] someone'. We have chosen to use the DOBJ function in these cases also, because nominal complements to adjectives have features in common with direct objects to verbs, but also because we want to avoid introducing separate functions with very limited scope. *Nærmest brudeparet* in a sentence such as (3.54) therefore gets the analysis in (3.55):

- (3.54) *Nærmest brudeparet beholdt kong Harald og Sven O.*
 nearest bridal couple+the kept King Harald and Sven O.
Høiby sine borddamer fra prosesjonen [...]
 Høiby their female dinner partners from procession+the
 'Closest to the bridal couple, King Harald and Sven O. Høiby kept their
 female dinner partners from the procession.' (Aftenposten 001)

(3.55)



3.4.2 Indirect Objects (IOBJ)

The function IOBJ is only used in structures where there is also a DOBJ-dependent¹¹. The indirect object's position in the sentence is before the direct object (Faarlund et al., 1997, 721) (unless the direct object is moved to the position before the finite verb, of course). In NDT, only nominal constituents receive the function IOBJ. Often, prepositional phrases headed by *til*, 'to', or *for*, 'for', have a semantics which is similar to an indirect object's semantics, but in our analyzes, such prepositions should have the function ADV (see 3.5).

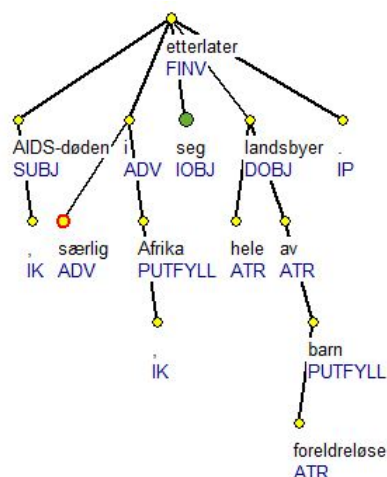
The highlighted words in the examples below, we analyze as IOBJ:

- (3.56) (a) *Han har gitt militsen ordre om å omgruppere,*
 he has given militia+the orders about to regroup
og kjempe.
 and fight
 'He has given the militia orders to regroup, and fight.' (Aftenposten 001)
- (b) *AIDS-døden, særlig i Afrika, etterlater seg hele*
 AIDS-death+the especially in Afrika leave behind refl entire
landsbyer av foreldreløse barn.
 villages of orphan children
 'The AIDS-death, especially in Africa, leaves behind entire villages of
 orphans.' (Aftenposten 001)

¹¹The exception is relative clauses where the direct object is relativized. In an example such as *Dette er boka som Per ga Kari*, 'This is the book which Per gave Kari', *Kari* will be IOBJ.

In (3.56 a), the indirect object is a noun, in (3.56 b), the reflexive pronoun *seg* receives the IOBJ function. The figure below shows the structure of (3.56 b):

(3.57)



3.4.3 Perception verbs and *be*, *la* etc. with object and infinitive

In NDT, we have chosen not to let clausal arguments receive a separate type of function. This has consequences for the analysis of structures containing an object/object-like phrase and an infinitive, often referred to as *accusative-with-infinitive*, as in the example below:

- (3.58) *Dersom du var på premieren, så du kanskje den lange mannen med den høye panna snike seg ut av kinosalen før filmen begynte.*
 if you were on premiere+the saw you maybe the long man+the with the high forehead sneak refl out of cinema+the before film+the started
 'If you were at the premiere, maybe you saw the long man with the high forehead sneak out of the cinema before the film started.' (Dagbladet 001)

The relevant constructions can be divided into two subgroups which we will give somewhat different analyzes, depending on whether the object-like nominal constituent receives a semantic role from the superordinate verb or not.

3.4.3.1 Perception verbs: “objects” as subjects for the infinitive

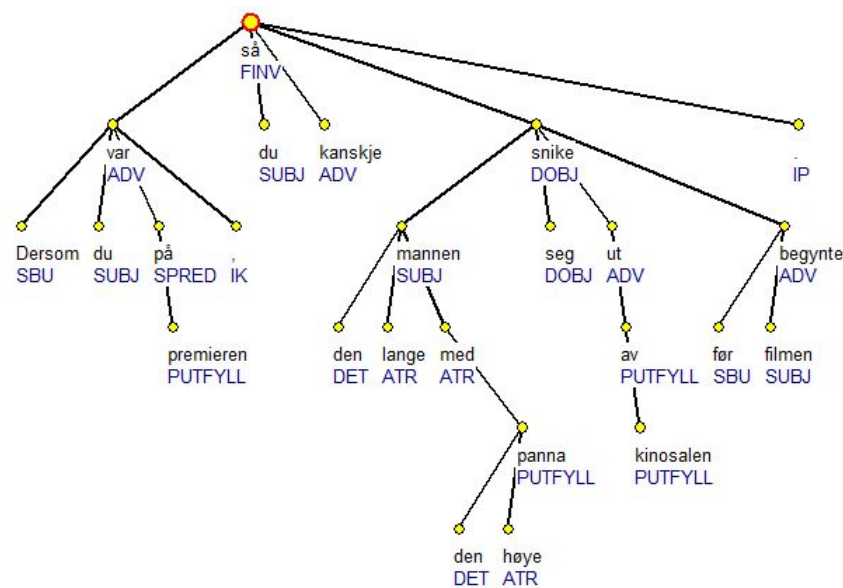
We will assume that perception verbs such as *høre*, *se* etc. do not assign any semantic role to the object-like nominal constituent. This is based on the possibility of sentences such as *Jeg har sett gass drepe*, 'I have seen gas kill'. As

gas cannot be seen, it is unlikely that *gass* is a semantic argument of *sett* (cf. Lødrup, 2008).¹²

We will therefore analyze the infinitive as DOBJ, while the object-like nominal constituent is dependent on the infinitive with the function SUBJ. There is therefore no direct dependency relation between the perception verb and the object-like nominal constituent. Example (3.58), repeated here, therefore receives the analysis in (3.59 b), where *den lange mannen...* is dependent on *snike*.

- (3.59) (a) *Dersom du var på premieren, så du kanskje den lange mannen med den høye panna snike seg ut av kinosalen før filmen begynte.*
 if you were on premiere+the saw you maybe the long man+the with the high forehead sneak refl out of kinosalen før filmen begynte.
 cinema+the before film+the started
 'If you were at the premiere, maybe you saw the long man with the high forehead sneak out of the cinema before the film started.'
 (Dagbladet 001)

(b)



3.4.3.2 *Be, la* etc.: IOBJ + DOBJ

Another group of verbs that can take an object-like nominal argument together with an infinitive, is *be*, 'pray/ask' and *la*, 'let', and other verbs with a similar meaning (*tillate*, 'permit', *påby*, 'order', etc.). Below we see an example:

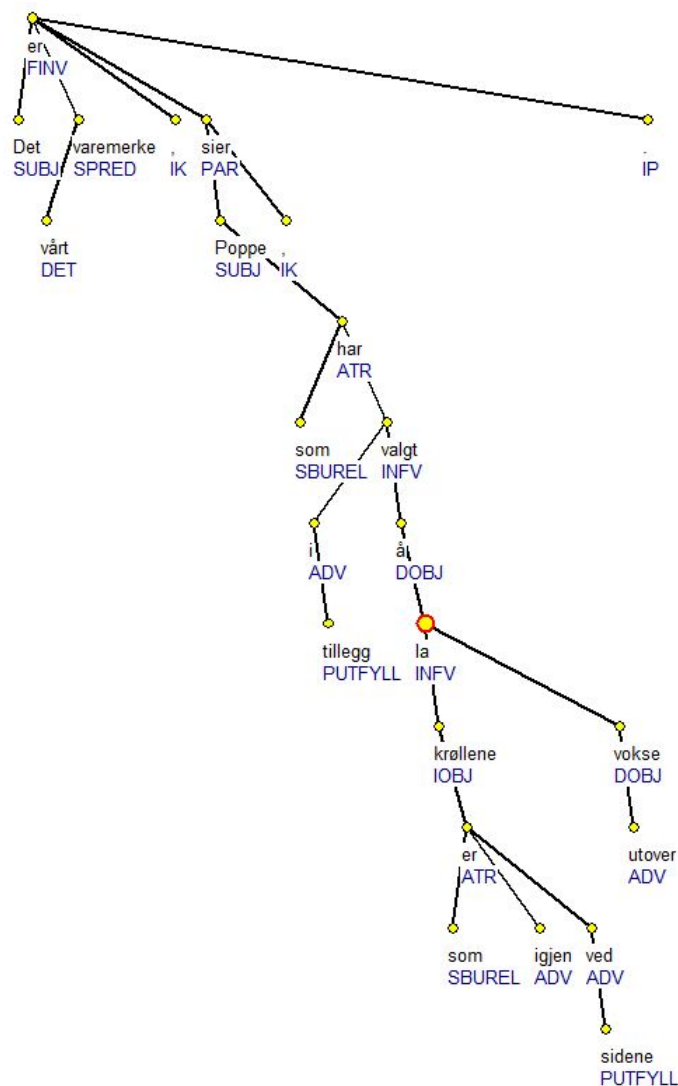
¹²Lødrup (2008) argues that such constructions are ambiguous, and our analysis will therefore not always be correct, according to that article. We have chosen not to pay attention to this ambiguity, to ensure a consistent annotation.

- (3.60) *Det er vårt varemerke, sier Poppe, som i tillegg har*
 it is our trademark says Poppe who in addition has
valgt å la krøllene som er igjen ved sidene vokse
 chosen to let curls which are left at sides+the grow
 utover.
 outwards
 'It is our trademark, Poppe says, who also has chosen to let the curls
 which are left at the sides, grow outwards.' (Dagbladet 001)

Here we assume that the object-like constituent (*krøllene...* in the above example) gets a semantic role from *be*, *la* etc. The verbs therefore have the object as a direct dependent, unlike the perception verbs in 3.4.3.1 above. (In many cases with *be* and *la*, we have object control, so that the object controls (is identical to) the subject of the infinitive. This is the case in the example above: *Krøllene* is identical to the subject of *vokse*.)

We will analyze the infinitive as DOBJ, just as we do with the perception verbs in 3.4.3.1. Since we have no other argument functions that may be suitable for the nominal constituent, and because we want to avoid introducing functions with a very limited use, we give the nominal object the function IOBJ. Example (3.60) above therefore gets the following analysis:

(3.61)



3.5 Adverbials (ADV)

In NDT, a large and diverse group of dependents get the function ADV. One reason for this is that we have chosen not to distinguish between obligatory and non-obligatory adverbials - all these are ADVs. This has, among other things, consequences for the analysis of prepositional phrases. Verbs such as *bo*, 'live [somewhere]', or *lete*, 'search', has an argument structure requiring a prepositional phrases as complement:

- (3.62) (a) *Jeg bor på landet.*
 I live on countryside+the
 'I live in the country.'

- (b) *Jeg leter etter kjolen.*
 I search after dress+the
 'I search for the dress'

The prepositions (*på* and *etter*) are in our analyzes ADV, just as prepositions in prepositional phrases with a looser association with the verb becomes ADV, for example, *i ti år* in the sentence below:

- (3.63) *Jeg har bodd på landet i ti år.*
 I have lived on countryside+the in ten years
 'I have lived in the country for ten years.'

It had been more linguistically adequate to distinguish between obligatory and non-obligatory adverbials. But it would probably be very difficult and time-consuming to implement such a distinction consistently in practice. In a corpus project like ours, where consistent analyzes are important and the time-frames are limited, it seems therefore prudent to use the same function for all types of adverbials.

Below we show how ADV function is used for dependents on verbs, adjectives, determiners, adverbs and prepositional phrases, and in special cases also on nouns. See Faarlund et al. (1997, 723-827) for a general overview of adverbials in Norwegian.

3.5.1 ADV to verbs

The ADV-function is used for dependents which modify the verb in different ways. This includes sentence adverbials, which modify the whole sentence. The emphasized constituents in the examples below are analyzed as ADV:

- (3.64) (a) *Kamskjell, piggvar og lammefilet sto på menyen*
 scallop turbot and fillet of lamb stood on menu+the
under den kongelige gallamiddagen.
 under the royal gala dinner+the
 'Scallop, turbot and fillet of lamb were listed on the menu at the royal gala dinner.' (Aftenposten 001)
- (b) *Men så er ikke bryllupet mellom Haakon og*
 but then is not wedding+the between Haakon and
Mette-Marit hverdagskost det heller.
 Mette-Marit daily fare that either.
 'But then, the wedding between Haakon and Mette-Marit isn't an everyday occurrence, either.' (Aftenposten 001)
- (c) *Ap. får styringen med de store internasjonale*
 Labour Party gets control+the with the big international
spørsmål på Stortinget, mens Høyre
 questions on Parliament while Conservative Party
konsentrerer seg om de nære ting.
 concentrate refl about the close things
 'The Labour Party gets the control over the big international questions in Parliament, while the Conservative Party concentrates on the close things.' (Aftenposten 001)

(d)

Andre ganger *strekker de seg så langt de kan, ved å*
 other times stretch they refl so far they can by to
love en ny vurdering av saken.
 promise a new evaluation og case+the
 'Other times, they go as far as they can, by promising a new
 evaluation of the case.' (Aftenposten 001)

(e)

Dette vervet bekles tradisjonelt av en representant
 this office occupy.pass traditionally by a representative
for opposisjonen.
 for opposition+the
 'This office is traditionally occupied by a representative of the
 opposition.' (Aftenposten 001)

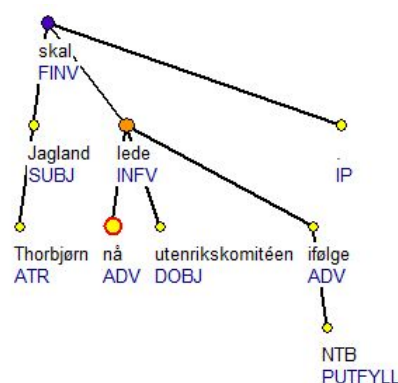
The examples above show some of the many forms ADVs can have. They can be prepositional phrases, as *sto på menyen* and *under den kongelige gallamiddagen* in (3.64 a), they can be adverbs, as in (3.64 b), subordinate clauses, as in (3.64 c), nouns, as in (3.64 d) or adjectives, as in (3.64 e).

In structures with both lexical verb and one or more auxiliary verbs, ADVs should be dependents on the lexical verb. This is illustrated below, where we see that the ADVs *nå* and *ifølge* are dependents on *lede*:

(3.65) (a)

Thorbjørn Jagland skal nå lede
 Thorbjørn Jagland will now lead
utenrikskomitéen ifølge NTB.
 Foreign Affairs Committee+the according to NTB
 'Thorbjørn Jagland will now lead the Foreign Affairs Committee,
 according to NTB.' (Aftenposten 001)

(b)



3.5.2 ADV to non-verbal heads: adjectives, determiners, adverbs and prepositions

Also adjectives, determiners, adverbs and prepositional phrases are modified by ADV-dependents. The modified constituents are then often themselves adjuncts (ATR, DET or ADV), either on a verb or a nominal constituent.

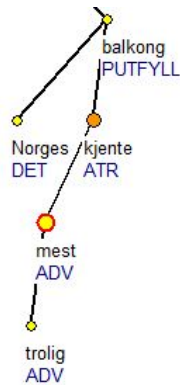
13

In some cases, the constructions are comparative constructions. See 5.8 for a detailed description of these. Note also that some adjectives sometimes can take DOBJs - this is explained in 3.4.1.2.

As with ADVs to verbs, ADVs to adjuncts may take many forms. The examples and figures below illustrate this (although this is not intended to be an exhaustive list).

- (3.66) (a) *Og tradisjonen tro kom prisvinnerne ut på Norges trolig mest kjente balkong.*
 and tradition+the faithful came price winners+the out on Norway's probably most known balcony
 'And, faithful to tradition, the price winners came out on Norway's probably best known balcony.' (Aftenposten 001)

(b)

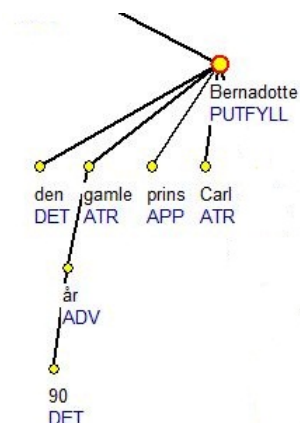


In (3.66 a), the adjective *mest* is ADV on *kjente*, while the adjective *trolig* is ADV on *mest*. In (3.67 a) below is a noun, *år*, which is ADV on the adjective *gamle*:

- (3.67) (a) [...] *mens dronning Silvia hadde overtatt sin "onkel Mulle", den 90 år gamle prins Carl Bernadotte.*
 while Queen Silvia had taken over her Uncle Mulle, the 90 years old prince Carl Bernadotte
 'while Queen Silvia had taken over her "Uncle Mulle", the 90 year old prince Carl Bernadotte.' (Aftenposten 001)

(b)

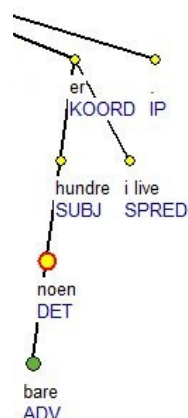
¹³We saw in 3.5.1 above that nouns sometimes can be ADV-dependents on verbs. We will not let such adjoint nouns get ADV-dependents themselves, however, but stick to the general rule that the noun's dependents are ATR or DET. In an example such as *allerede mandag*, 'already Monday', therefore, *allerede* will be ATR on *mandag*. The only times nouns should have ADV-dependents, is the elliptical constructions, see 3.5.3 below.



In (3.68 a) below, it is not an adjective, but a determiner, *noen*, which is modified. We see that *bare* has gotten the function ADV:

- (3.68) (a) [...] *bare noen hundre er i live.*
 only some hundred are alive
 'Only a few hundred are alive' (Dagbladet 001)

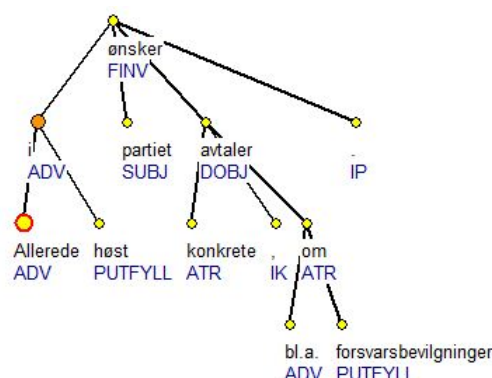
(b)



Also prepositions can be modified by ADV-dependents. In the example below we see that the adverb *allerede* is ADV on *i*. We also see that *bl.a.*, which is a preposition in our analyses, is ADV on *om*:

- (3.69) (a) *Allerede i høst ønsker partiet konkrete avtaler,*
 already in fall wished party+the concrete agreements,
 bl.a. om forsvarsbevilgninger.
 e.g. about defence fundings
 'Already this fall, the party wished concrete agreements about
 defence fundings' (Aftenposten 001)

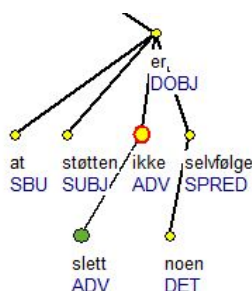
(b)



Not surprisingly, adverbs can be modified by ADV-dependents. Below we see that *ikke* has *slett* as ADV-dependent:

- (3.70) (a) [...] *at støtten slett ikke er noen selvfølge.*
 that support+the simply not is any matter of course
 '... that the support simply shouldn't be taken for granted'
 (Aftenposten 001)

(b)

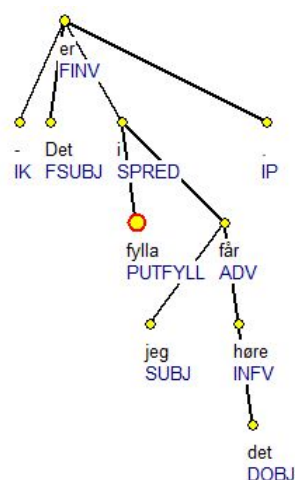


We will sometimes find also subordinate clauses/infinitives that are ADVs to words with non-verbal functions. Two types should be mentioned specifically: firstly constructions with clefting where an adjective, adverb or prepositional phrases is relativized (see 3.3 for more on clefting. See also 3.10.3). Here, the relative-like clause becomes an ADV-dependent on the relativized adjective/adverb /prepositional phrase. In the example below, is *jeg får høre det* therefore ADV on the head of the prepositional phrase *i fylla*.¹⁴

- (3.71) (a) - *Det er i fylla jeg får høre det.*
 it is in drunkenness I get hear it
 'It is when drinking I hear it.' (Dagbladet 001)

(b)

¹⁴Note that the relativized constituent is not an adjunct, unlike in other constructions we have described in this subsection. *I fylla* has the function SPRED.



The second construction we will mention, is adjectives that occur with infinitives, as in the example below:

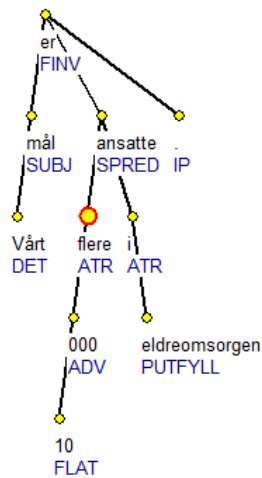
- (3.72) *Det finnes ikke mange steder som er gode å være når*
 it exists not many places which are good to be when
man opplever noe slikt.
 one experiences something this kind
 'There are not many places where it is good to be when you experience
 something like that.' (Dagbladet 003)

Here, *å være* is analyzed as ADV on *gode*. See 3.7 for more on this construction.

In addition to adjectives, adverbs, determiners and prepositions, ADVs may sometimes also modify complementizers. See 3.10 for a description of this.

In constructions such as *10 000 flere ansatte*, '10 000 more employees', the numeral is not used to determine the number of employees directly, but is used instead to modify the adjective *flere* to show rise in number of employees. This fact is reflected in the syntactic annotation, and *10 000* is dependent on *flere*, and takes the function ADV. *Flere* gets a standard analysis, i.e. it is ATR on *ansatte*.

- (3.73) (a) *Vårt mål er 10 000 flere ansatte i eldreomsorgen.*
 our goal is 10 000 more employees in geriatric care+the
 'Our goal is 10 000 more employees in the geriatric care.' (Stortinget
 005)
 (b)



Distributive use of numerals means that a numeral indicates an amount equally distributed among each member of a group. (3.74) is an example:

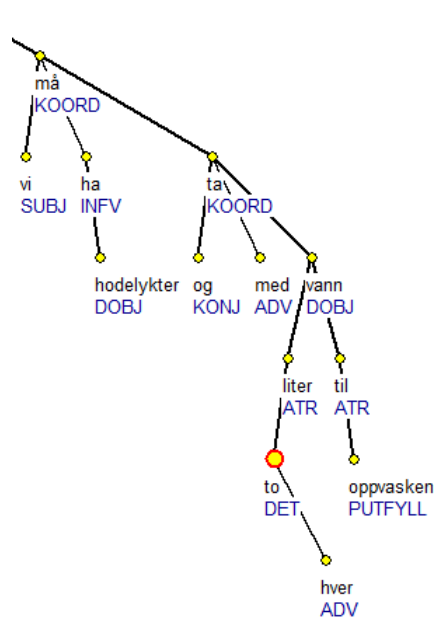
- (3.74) [...] *vi må ta med hodelykter og to liter vann*
 we must take with headlamps and two liters water
hver til oppvasken.
 each to washing-up+the
 'We must take with us headlamps and two liters of water each for the
 washing-up.' (Dagbladet 014)

Hver functions as a marker of distributivity, and it may be tempting to analyze it as dependent on *vann*, but note that there is no concord (**to liter vann hvert*). *Liter* is masculine, but if we replace with a neuter word, there is still no concord (**to glass vann hvert*). A paraphrase shows that *hver* refers to the people in the group (*hver person tok med to liter vann*, 'each person brought two liters of water'), but note that even here there is no actual concord. We can see this if the group consists of neuter objects, such as *rom* in the following example:

- (3.75) **Alle rommene i huset har to vinduer hvert*
 all rooms+the in house+the has two windows each.neut
 'All rooms in the house have two windows each.'

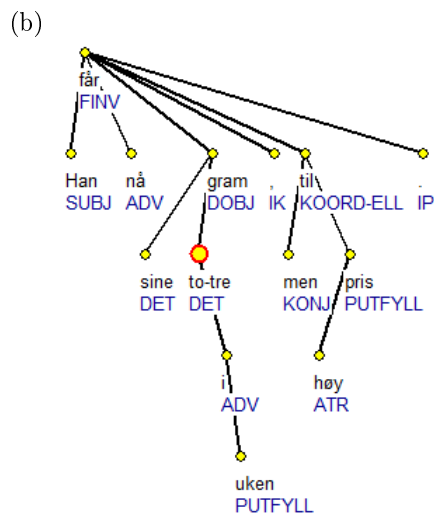
Our conclusion is that *hver* in such distributive constructions does not have any actual determiner use on a noun, as this word would otherwise have had. Rather, it occurs noninflected as a particle, marking that the numeral should be understood distributively. We have therefore chosen to analyze it as a dependent on the numeral with the function ADV:

- (3.76)



There is another type of distributive construction, where we, instead of *hver* find a time adverbial, but where the syntactic and semantic relationships are fairly parallel. Consequently, we use the same analysis here, and put the time adverbial as an ADV-dependent on the numeral.

- (3.77) (a) *Han får nå sine to-tre gram i uken, men til en høy pris.*
 he gets now his two-tre grams in week+the but to
 a high price
 'He now gets his two grams a week, but to a high price.'
 (Aftenposten 011)



3.5.3 ADV to nouns

As stated in the paragraphs above, we nouns usually cannot take ADV-dependents in NDT (see especially 13). The exceptions are constructions where we assume that a verb is omitted, i.e. elliptical constructions. In the example below, *aldri* and *uten at passet er med* are ADV-dependents on *dag*:

- (3.78) *Aldri en dag uten at passet er med.*
 never a day without that passport+the is with
 'Never a day without bringing the passport' (Dagbladet 001)

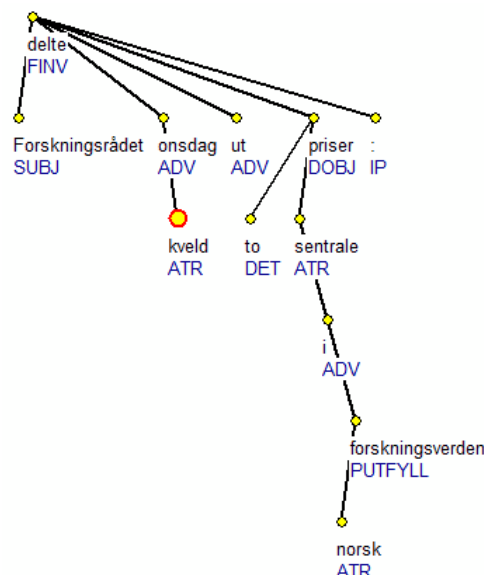
See 5.1.2 for a description and illustration of this.

3.5.4 Complex time adverbials

Time adverbials such as *tirsdag kveld*, 'Tuesday evening', and *i går ettermiddag*, 'yesterday afternoon', are analyzed with *kveld* and *ettermiddag* as ATR-dependents on *tirsdag* and *i går*, respectively. The principle here is that the last element in the complex time adverbial is dependent on the first.

- (3.79) (a) *Forskningsrådet delte onsdag kveld ut to sentrale priser i norsk forskningsverden:*
 Research Council handed Wednesday evening out two central prizes in Norwegian research world+the
 'The Research Council handed out two central prizes in the Norwegian research world Wednesday evening:' (Aftenposten 007)

(b)



3.5.5 Path-constructions

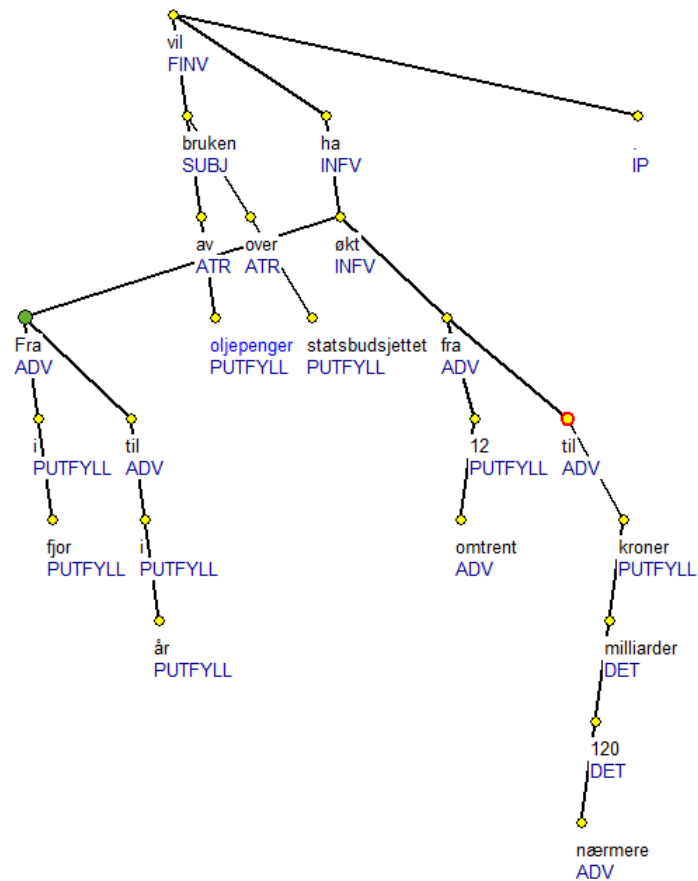
Where two prepositions combine to describe a route for a movement, i.e. a structure of the type *from X to Y* (called a path-construction), we let the *to*-phrase be dependent on *from*, with the function ADV. Note that these routes

may be metaphorical or describe a period of time. Examples of both are found in (3.80 a):

(3.80) (a)

Fra i fjor til i år vil bruken av oljepenger
 from in last year to in year will use+the of oil money
 over statsbudsjettet ha økt fra omtrent 12
 over state budget+the have increased from approximately 12
 til nærmere 120 milliarder kroner.
 to near 120 billion crowns
 'From last year to this year, the spending of oil money on the state
 budget will have increased from approximately 12 to close to 120
 billion crowns.' (Aftenposten 007)

(b)



In more complex path constructions, such as *from X through Y to Z*, *from* will still be the head, and each subsequent preposition is dependent with ADV on the immediately preceding preposition, so that *through* becomes ADV on *from*, and *to* becomes ADV on *through*.

3.6 How to distinguish between transitive and intransitive prepositions

Prepositions can be both transitive and intransitive, i.e. with and without a complement (PUTFYLL in our annotation). In some sentences, it may be uncertain whether an “object-like” phrase is the object of a verb or the complement of a preposition.

In the examples below, it is not necessarily obvious if *en sang eller film ...*, *de store tingene* and *den nye boka* is dependent on the verb or the preposition.¹⁵

- (3.81) (a) [...] *Setter på en sang eller en film som passer til humøret hennes.*
 puts on a song or a film which fits to mood+the her
 ‘[She] puts on a song which fits to her mood.’ (Dagbladet 001)
- (b) *Vi merket veldig fort at vi begynte å snakke om de store tingene.*
 we realized very quickly that we started to talk about the big things+the
 ‘We realized very quickly that we started to talk about the big questions.’ (Dagbladet 001)
- (c) *Ho leitte etter den nye boka.*
 she looked after the new book+the
 ‘She looked for the new book.’ (Faarlund et al., 1997, 697)

Below we present three tests that can be helpful in such cases. The tests are based on Faarlund et al. (1997, 699-700).

1. The position of the preposition

In constructions with intransitive prepositions, an object that has the form of an unstressed pronoun noun can (and must), move to the position before the preposition. A pronoun which is a complement to transitive the preposition, may not, however, move in this way:

- (3.82) (a) *Han satte den på.*
 He put it on
 ‘He put it on.’
- (b) **Ho leitte den etter.*
 she looked it after
 ‘*She looked it after.’

We will analyze (3.82 a) with intransitive preposition and (3.82 b) with transitive preposition.

¹⁵Some intransitive prepositions, as in sentences like *Han slo på lyset*, ‘He turned on the light’, are often referred to as verbal particles. In an initial phase of the project, we tried to annotate these with a separate syntactic function, but this proved very difficult to implement in practice, so we have gone away from it.

2. Definiteness in impersonal passive sentences

Norwegian has the restriction that an NP which is potential subject in a presentational sentence must be indefinite (see 3.2.6). The same applies to objects in the impersonal passive sentences (Faarlund et al., 1997, 846). The restriction does not apply to NPs which are complements to a preposition. We can therefore turn a sentence into an impersonal passive, and test whether the object/prepositional complement can be definite:

- (3.83) (a) **Det ble satt på sangen.*
 it became put on song+the
 'The song was put on'
 (b) *Det ble lett etter den nye boka.*
 it became looked after the new book+the
 'The new book was looked for.'

Example (3.83 a) does not work, and we can therefore assume that *sangen* is the object of the verb, and not complement of *på*. Example (3.83 b) works fine, however, suggesting that *den nye boka* is a prepositional complement, not the object of the verb.

3. Deletion of the preposition in coordination

If the same verb occurs in two coordinated sentences with different subject and object, the verb may be deleted in the second sentence. A preposition with a complement must be repeated in the second sentence, while an intransitive preposition can be deleted. In the first of the two examples below, the preposition is deleted, which suggests that it is an intransitive preposition. In the second example, deletion doesn't work as well, suggesting that *boka* and *katten* are prepositional complements.

- (3.84) (a) *Per satte på musikken og Kari filmen.*
 Per put on music+the and Kari film+the
 'Per put on the music and Kari the film.'
 (b) *??Per lette etter boka og Kari katten.*
 Per looks after book+the and Kari cat+the
 '*Per looks for the book and Kari the cat.'

Conversely, the first of the examples below doesn't work without deletion of the preposition, which is an argument for analyzing it as an intransitive preposition. In the second example, the preposition can stay, and we therefore count it as transitive.

- (3.85) (a) **Per satte på musikken og Kari på filmen.*
 Per put on music+the and Kari on film+the
 '*Per put on the music and Kari on the film.'
 (b) *Per lette etter boka og Kari etter katten.*
 Per looked after book+the and Kari after cat+the
 'Per looked for the book and Kari for the cat.'

Sometimes it's not a nominal constituent that creates doubt, as in the examples we discussed above, but a clause (or an infinitive). In the example below, it is

not obvious whether the clause *at familiemedlemmene...* should be complement to *med* or not.

- (3.86) *Nå regner hun med at familiemedlemmene har blitt*
 now expects she with that family members have become
så vant til arbeidet hennes at denne utstillingen ikke
 so used to work+the her that this exhibition+the not
skal bidra til ytterligere uro.
 shall contribute to further unrest
 ‘She expects that the family members have become so used to her work
 now that this exhibition will not contribute to further unrest.’ (Dagbladet
 003)

In most such cases, it is usually possible to replace the the clause/infinitive with a pronoun (*det*), and all the tests we have described above can therefore be used.

We sometimes also encounter prepositional phrases for which it may be unclear whether they are themselves prepositional complements or arguments of the verb (cf. 4.3). In the example below, it is not necessarily obvious whether *på dem* should be a complement of *ned* or *ser*:

- (3.87) *Jeg ser på ingen måte ned på dem [...].*
 I look on no way down on them
 ‘I don’t look down on them in any way.’ (Dagbladet 001)

Here, the tests in 1 and 2 do not work, since they are based on properties of NPs. Our experience is that the test in 3 does not give clear results either. The test assumes grammaticality judgments of examples such as (3.88 a) and (3.88 b) below, but here our intuitions are unclear and inconsistent.

- (3.88) (a) *??Kari ser ned på svenskene og Per ned på*
 Kari looks down on Swedes+the and Per down on
 danskene.
 Danes+the
 ‘??Kari looks down on Swedes and Per down on Danes.’
 (b) *??Kari ser ned på svenskene og Per på danskene.*
 Kari looks down on Swedes+the and Per on Danes+the
 ‘??Kari looks down on Swedes and Per on Danes.’

The analysis of examples such as (3.87) is therefore to some extent up to the judgment of the annotators. (In the specific example, the annotator has chosen to analyze *ned* as a transitive preposition.)

3.7 Non-finite verbs

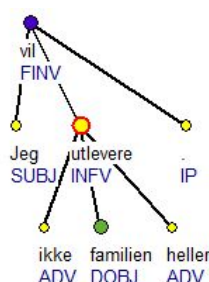
In this section we present the analysis of different constructions with non-finite verb forms. 3.7.1 discusses the relationship between auxiliary verbs and lexical verbs, 3.7.2 describes the analysis of infinitival clauses with an infinitival marker, while 3.7.3 discusses infinitival clauses without an infinitival marker. 3.7.4 is devoted to the analysis of participles.

3.7.1 Auxiliary verbs and lexical verbs

In constructions with auxiliary verbs, the auxiliary should be the head and the lexical verb be dependent with the function INFV. Sentence (3.89 a) should have the analysis in (3.89 b):

- (3.89) (a) *Jeg vil ikke utlevere familien heller.*
 I will not expose family+the after all
 'After all, I don't want to expose my family.' (Dagbladet 003)

(b)



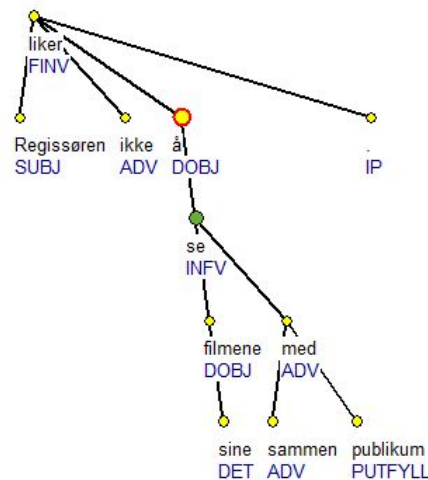
For more on this, see 3.1.4.

3.7.2 Infinitival clauses with an infinitival marker

In the case of infinitival clauses with an infinitival marker, as in (3.90 a), we have chosen to let the infinitival marker be the head and let the infinitive be dependent on the infinitival marker with the function INFV. The reason we have chosen this solution, is that we want an analysis that is parallel to the one we have chosen for infinitives with auxiliary verbs. The infinitival marker carries the grammatical function and the infinitival clause. In (3.90 a), the infinitival marker will therefore get the function DOBJ, with the infinitive *se* as INFV-dependent, cf. (3.90 b):

- (3.90) (a) *Regissøren liker ikke å se filmene sine sammen med publikum.*
 director+the likes not to see films+the his together
 with audience
 'The director doesn't like to see his films together with the audience.'
 (Dagbladet 001)

(b)



Infinitival clauses with an infinitival marker can have a variety of grammatical functions in the sentence in addition to direct object, such as subjects, as in (3.91 a), predicative, as in (3.91 b) and potential subjects, as in (3.91 c) (for potential subjects in such constructions, see 3.2.3). They can also be prepositional complements, as in (3.91 d):

- (3.91) (a) *Å spise fisk er sunt.*
 to eat fish is healthy
 'To eat fish is healthy.'
- (b) *Den største utfordringen er å legge vekk farsrollen.*
 the biggest challenge is to put away father role+the
 'The biggest challenge is to put away the role as a father.'
 (Dagbladet 003)
- (c) *Det er viktig å lage noe som er spesielt*
 it is important to make something which is special
 [...].
 'TIt is important to make something which is special.' (Dagbladet 002)
- (d) *En av de innsatte på Ila sitter de for å ha drept*
 one of the inmates on Ila sits there for to have killed
 tre mennesker.
 three people
 'One of the inmates at Ila is there because of having killed three
 people.' (Dagbladet 001)

Infinitival clauses with an infinitival marker can also be dependents on adjectives, as in (3.92 a), and they can be attributive relative sentences, as in (3.92 b).

- (3.92) (a) *Det finnes ikke mange steder som er gode å være*
 it exists not many places which are good to be
 når man opplever noe slikt.
 when one experiences something this kind

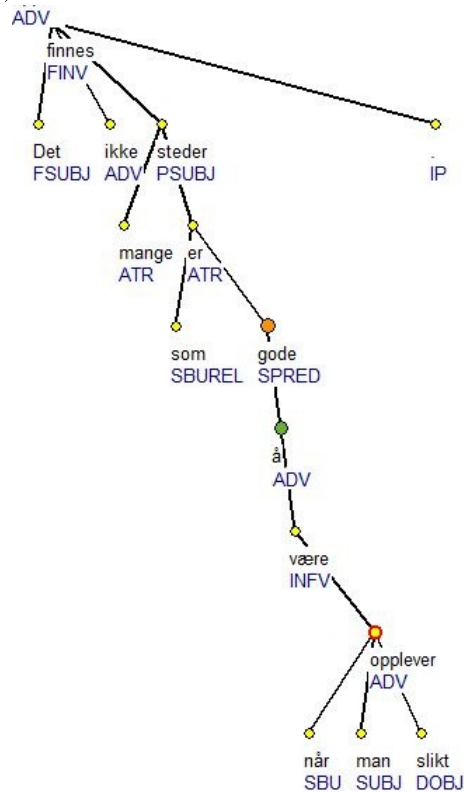
'There are not many places where it is good to be when you experience something like that.' (Dagbladet 003)

- (b) *Det er jo alltid en melodi å lage.*
 it is certainly always a melody to make

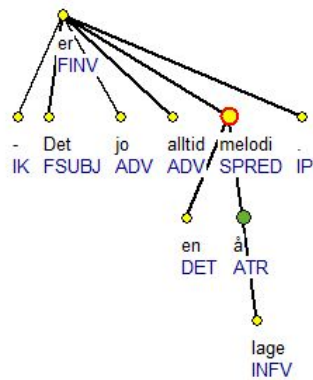
'There is always a melody to make.' (Dagbladet 003)

Infinitives to adjectives should have the function ADV (cf. Faarlund et al. 1997, 406), while infinitival relatives should be ATR. The analyzes of (3.92 a) and (3.92 b) is therefore as follows:

(3.93) (a)



(b)

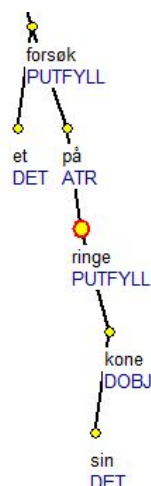


3.7.3 Infinitival clauses without an infinitive marker

Some verbs take infinitival objects without an infinitive marker, such as perception verbs and *be* and *la*. In such cases, the infinitive gets the function direct object (DOBJ). For more on these constructions, see 3.4.3. Sometimes infinitives occur without an infinitival marker where one would normally expect an infinitival marker, as in (3.94 a). In such cases, the infinitive gets the function the infinitive marker should have had. In (3.94 a) the function of the infinitive is therefore PUTFYLL, cf. (3.94 b):

- (3.94) (a) [...] *et forsøk på ringe sin kone.*
 an attempt on call his wife
 ‘an attempt to call his wife.’ (Dagbladet 003)

(b)



3.7.4 Participles

Bokmål In the morphological annotation for Bokmål, a participle can either have *verb* or *adjective* as part of speech. Participial attributes which precede their head noun shall normally be adjectives. In other cases, annotators have to choose between the two parts of speech. For more on this, see 1.2.1.

In the case of participles to *være* and *bli*, as in (3.95 a) and 3.95 b), the choice of part of speech has implications for the syntactic analysis.

- (3.95) (a) - *Jeg er litt plaget av revmatisme [...].*
 I am a little bothered by rheumatism

‘I am a little bothered by rheumatism.’ (Dagbladet 003)

- (b) *Erik Poppe blir funnet livløs i senga.*

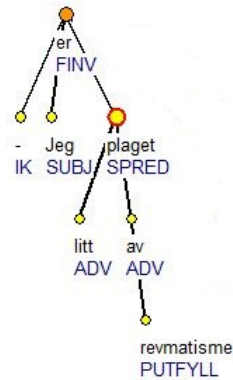
Erik Poppe becomes found lifeless in bed+the

‘Erik Poppe is found lifeless in the bed.’ (Dagbladet 001)

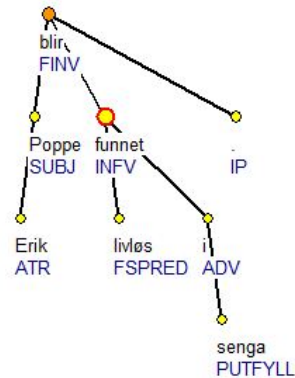
If the participle has the part of speech *adjective*, we will normally give it the function SPRED. This is the analysis we have chosen for (3.95 a), as (3.96 a)

shows. If the participle has the part of speech *verb*, it will be natural to understand it as part of a periphrastic passive construction. The participle will therefore get INFV. This is the analysis we have chosen (3.95 b), as (3.96 b) shows.

(3.96) (a)



(b)



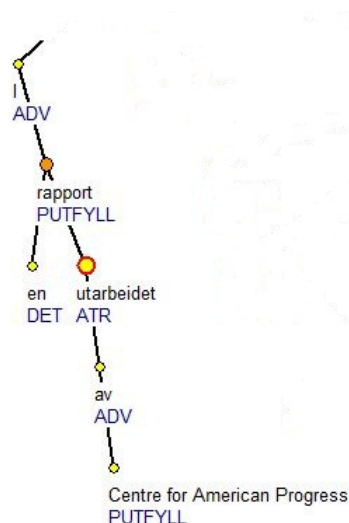
Sometimes we find participles after nouns, as in (3.97).

(3.97) *I en rapport utarbeidet av Centre for American Progress*
 In a report prepared of Centre for American Progress
 [...].

'In a report prepared by the Centre for American Progress'
 (Klassekampen 001)

They meet the criteria for verbal participles (cf. 1.2.1). At the same time there are good reasons to count them as modifiers of nouns, and they can often be replaced by a relative clause without the meaning changing. We analyze such participles as attributes, see (3.98).

(3.98)



Nynorsk For technical reasons, all perfect participles in Nynorsk after *vere* and *verte/bli* will unfortunately be tagged as *adjective*, as we have explained in 1.2.1.2. In the syntax, we will still distinguish between adjectival and verbal participles in this environment, giving the former the function SPRED and the latter INFV.

Attributive participles which follow their head noun of the kind we saw in (3.97), will not have the part of speech *verb* in the morphological annotation either.

Present participle Present participles are always adjectives in the morphological annotation, and will be treated as adjectives in the syntax. Present participle with *bli*, as in (3.99), will therefore be regarded as SPRED, not INFV.

(3.99) *Han ble sittende.*
 he became sitting
 'he remained seated'

3.8 Predicative

A predicative is characterized by a kind of double dependency: On the one hand it is clearly a dependent of the verb. On the other hand, it characterizes or identifies another word in the sentence, usually the subject or the object. In (3.100a), the predicative, *syk*, is a property of the subject, *Per*. In (3.100b), it is the object, *huset*, which is characterized by the predicative, *rødt*.

(3.100) (a) *Per er syk.*
 Per is ill
 'Per is ill.'

- (b) *De malte huset rødt.*
 they painted house+the red
 'They painted the house red.'

As we accounted for in 2.2, we do not use multiple dependencies on the same node in NDT. Which word the predicative characterizes or identifies, will therefore follow from the syntactic function, not by the dependency structure: In (3.100a) the function will be SPRED (subject predicative), while in (3.100b), it will be OPRED (object predicative).

Predicatives can furthermore be either part of the verb's argument structure or free phrases. In (3.100a) and 3.100b), the predicatives are arguments of the verb, they can not be removed without changing the meaning completely. In (3.101a) and (3.101b), however, the predicatives are not required by the verb. The predicative *rå* in (3.101a) describes how the object, *gulrota*, was when Per ate it. It can be removed without the verb meaning changing. In a similar way, *skjelvende* in (3.101b) is a characteristic of the subject, *Per*, at the time he closed the door. This is not required by the verb either.

- (3.101) (a) *Per spiste gulrota rå.*
 Per ate carrot+the raw
 'Per ate the.'
 (b) *Per lukket døra, skjelvende i panikk.*
 Per closed dor+the shivering in panic
 'Per closed the dor, shivering with panic.'

Such free predicatives get the functions FSUBJ (free subject predicative) and FOPRED (free object predicative). Free predicatives may in some cases refer back on other phrases than subjects and objects. In such cases we still use the function FOPRED.

In the use of the predicative functions we follow the chapter on predicative in NRG to a large extent (Faarlund et al. 1997, 732-773). In the following, we will not repeat what is said there, but come with some clarifications relevant for the annotation.

3.8.1 Copular verbs

The verbs *være* and *verte/bli* will, when they are not auxiliary verbs, normally have an SPRED-dependent. The exceptions are cases where *være* has a clear existential meaning and can be replaced by verbs such as *eksisterer* or *finnes*:

- (3.102) *Å være eller ikke være, det er spørsmålet.*
 to be or not be it is question+the
 'To be or not to be, that is the question'

In the case of presentational sentences, *være* does not take SPRED-dependents, see 3.3.

In Faarlund et al. 1997, 737-738 a distinction is made between locative prepositions after *være*, which is analyzed as an obligatory adverbial, and other prepositions, which are analyzed as subject predicative. We have chosen not to implement this distinction. The preposition in both (3.103a) and (3.103b) shall have SPRED:

- (3.103) (a) *Per er i byen.*
 Per is in city+the
 'Per is in the city.'
 (b) *Per er i god form.*
 Per is in good shape
 'Per is in good shape.'

3.8.2 OPRED with *ha*

In Faarlund et al. 1997, 752-753 it is mentioned that the verb *ha* can take a present participle as object predicative, as in (3.104):

- (3.104) *Jeg har en gammel portvin stående.*
 I have a old port wine standing
 'I have an old port wine in reserve.' Faarlund et al. 1997, 752

In our texts, *ha* appears with a number of other types of predicatives. In (3.105 a), we have an adjective, in (3.105 b), the preposition *som* (which is common with predicative nouns, cf. Faarlund et al. 1997, 754) and in (3.105 c), we have the preposition *på*. In such cases we use OPRED.

- (3.105) (a) *Det er greit å ha en kølle klar når den gamle*
 it is ok to have a baton ready when the old
brekker.
 breaks
 'It is ok to have a baton ready when the old one breaks.' (Dagbladet 001)
 (b) *Hvordan er det for sønnen din å ha far som sjef*
 how is it for son+the your to have dad as boss
i bandet?
 in band+the
 'How is it for your son to have his father as boss in the band?' (Dagbladet 003)
 (c) *De rødgrønne i Trondheim mener kulturlivet i*
 the red-green.pl in Trondheim think culture life+the in
byen har fått et kjempeløft mens de har hatt
 city+the has got a huge lift while they have had
hånda på rattet i styringa av
 hand+the on steering wheel+the in administration+the of
kulturbudsjettet.
 cultural budget+the
 'The Red-Green in Trondheim think that the culture life in the city has got a face lift' (Klassekampen 001)

Examples such as (3.105 c) can be difficult to distinguish from sentences with have direct objects and place adverbial, as in (3.106):

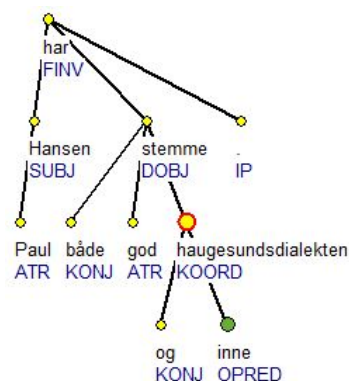
- (3.106) *Per hadde et raseriutbrudd i butikken.*
 Per had a temper tantrum in shop+the
 'Per had a temper tantrum in the shop.'

An important difference between (3.105 c) and (3.106) is that the prepositional phrase in (3.105 c) is required for determining the intended meaning: (3.105 c) does not mean that the coalition has a hand. In (3.106), we can delete the prepositional phrase without the meaning changing.

An object after *ha* can, interestingly enough, be coordinated with an object + a object predicative, as in (3.107a). When the object predicative is in the second conjunct, we let it be dependent on the coordinated object, cf. (3.107b)¹⁶:

- (3.107) (a) *Paul Hansen har både god stemme og*
 Paul Hansen has both good voice and
haugesundsdiakten inne.
 Haugesund dialect+the inside
 'Paul Hansen has both a good voice and masters the Haugesund
 dialect.' (Dagbladet 003)

(b)



3.8.3 Predicative after *med*

After *med*, 'with', (and probably also after *uten*, 'without'), you can have a predicative which is connected to the complement of the preposition, as in (3.108 a) and (3.108 b):

- (3.108) (a) *Prinsesse Märtha Louise savnet neppe sin Ari med*
 Princess Märtha Louise missed hardly her Ari with
statsminister Jens Stoltenberg som kavalier [...].
 Prime Minister Jens Stoltenberg as partner
 'Princess Märtha Louise hardly missed her Ari with Prime Minister
 Jens Stoltenberg as partner.' (Aftenposten 001)

¹⁶An alternative would be to consider this as sentence coordination with verbal ellipsis, cf. 5.1.3. We don't believe, however, that we are dealing with verbal ellipsis here. Firstly, such coordination is also possible in the related construction described in 3.8.3, where verbal ellipsis is not an option. Secondly, it doesn't appear to be possible to have verbal ellipsis with OPRED in only one conjunct: **Per malte hytta rød, og huset*, 'Per painted the cabin red and the house'.

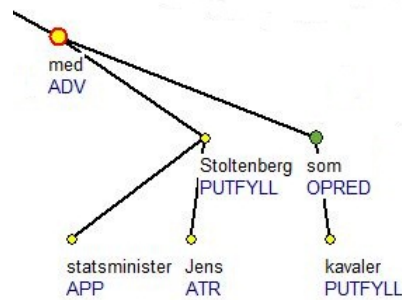
- (b) *Hver spiller har fast plass, men navn og nummer*
 every player has permanent space with name and number
skrevet inn på en plakett over seg.
 written inn on a plaque over refl
 'Every player has a permanent space, with his name and number
 written on a plaque over him.' (Dagbladet 001)

This phenomenon is described in Haug 2009 (but see also Faarlund et al. 1997, 458-459). The construction has a meaning that is quite similar to object predicative with *ha*, and it is often possible to paraphrase with this:

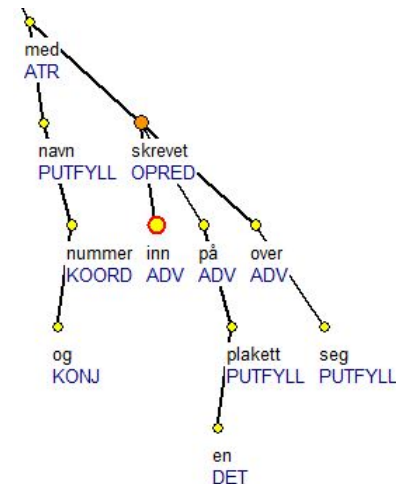
- (3.109) *Märtha Louise har Stoltenberg som kavalier.*
 Märtha Louise has Stoltenberg as partner
 'Märtha Louise has Stoltenberg as partner.'

We let the predicative be OPRED-dependent on the preposition, cf. (3.110 a) and (3.110 b):

- (3.110) (a)



- (b)

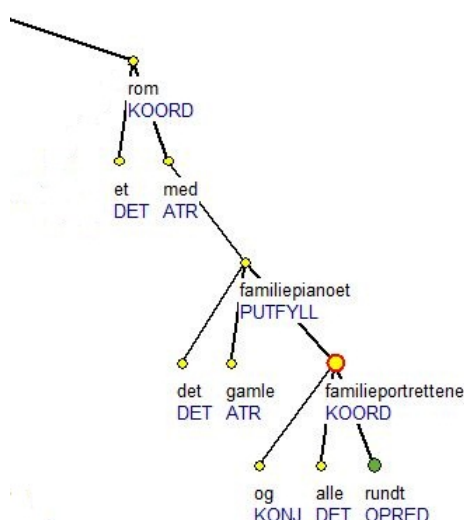


Also for predicative with *med*, coordination that corresponds to the one we saw in (3.107 a), occurs: A prepositional complement to *med* can be coordinated with a prepositional complement + a the predicative, cf. (3.111):

- (3.111) [...] *et rom med det gamle familiepianoet og alle familieportrettene rundt.*
 a room with the old family piano+the and all
 family portraits+the around
 'a room with the old family piano and all the family portraits around.'
 (Dagbladet 003)

In such cases, we let the predicative be dependent on the second conjunct:

(3.112)



3.8.4 FOPRED to non-objects

Free predicatives sometimes refer back to words that are not subject or object, as in (3.113).

- (3.113) *Som ordboksforfatter er det stadig like ergerlig [...]*
 as dictionary author is it constantly just as annoying
å se sine ting igjen på denne måten.
 to see ones things again on this way+the
 'As a dictionary author, it is always just as annoying to see ones own
 things again in this way.' Faarlund et al. 1997, 760

In the corpora, such predicatives get the function FOPRED.

3.9 Interjections

Words with part of speech *interjection* must have the function INTERJ. INTERJ can be a root function, cf. sentence (3.114 a) and the analysis in (3.114 b).

For more on the use of INTERJ as a root function, see 3.1.2.

- (3.114) (a) - *Ja.*
 yes
 'Yes.' (Dagbladet 003)

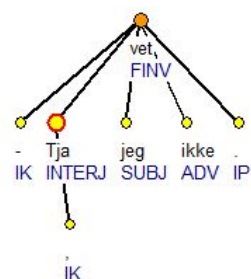
(b)



Interjections may also be part of a sentence. Then interjection will typically be dependent on a verb, see (3.115 a) and (3.115 b):

- (3.115) (a) - *Tja, jeg vet ikke.*
 yes and no I know not
 'Well, I don't know.' (Dagbladet 001)

(b)



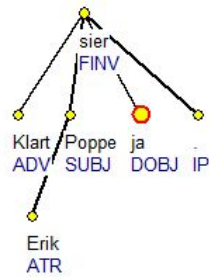
Note that *ja* and *nei* can have both *interjection* and *noun* as part of speech. The part of speech *interjection* and the function INTERJ will be used e.g. when the answer to a question is quoted directly, as in (3.116):

- (3.116) - *Er du blitt mer følsom med åra?* - *Ja.*
 are you become more sensitive with years+the yes
 '- Have you become more sensitive as the years pass? - Yes.' (Dagbladet 001)

However, *ja* and *nei* are often direct objects to verbs like *svare*, etc. The part of speech is then *noun* and the function is DOBJ:

- (3.117) (a) *Klart Erik Poppe sier ja.*
 obviously Erik Poppe says yes
 'Obviously, Erik Poppe says yes.' (Dagbladet 001)

(b)



3.10 Subordinate clauses

Here, we will look at the analysis of finite subordinate clauses, while non-finite subordinate clauses are described in detail in 3.7 and 3.4.3. In 3.10.1 we will present the fundamental choices of analysis, 3.10.2 describes the analysis of the complementizer. The analysis of relative clauses and indirect interrogative clauses are explained respectively in 3.10.3 and 3.10.4.

3.10.1 Fundamental choices of analysis

In NDT, the finite verb will always be the head for a subordinate clause. In some corpora, including PROIEL (Haug 2010, 36), the complementizer is the head. When we have chosen to let the finite verb be the head, it is partly because we want the analysis of subordinate clauses and matrix clauses to be as similar as possible (see 3.1.1). Moreover, it is advantageous to analyze subordinate clauses with and without complementizer the same way (cf. 3.10.2).

We have no separate functions for subordinate clauses, but use the same functions as for other constituents: The clause in (3.118 a) should for example be SUBJ, (3.118 b) DOBJ, (3.118 c) ADV, (3.118 d) PUTFYLL and (3.118 e) ATR (cf. 3.10.3).

(3.118) (a)

Om platene noen gang kommer for salg, *vites*
 about plates+the any time come for sale know.pass
ikke.
 not
 'If the plates ever come for sale, isn't known.' (Dagbladet 001)

(b) *Vi visste jo* at de ikke fyrte noe
 we knew certainly that they not have heating on any
særlig uansett.
 special anyway
 'We knew that they didn't have the heating on a lot anyway.'
 (Dagbladet 003)

(c)

Da hun var tenåring, *begynte hun å sy dem selv*.
 When she was teen stared she to sew them herself

'When she was a teen, she started to sew them herself' (Dagbladet 002)

- (d) *Jeg hadde en idé om at jeg skulle være dekket av*
 I have an idea about that I should be covered of
roser [...].
 roses

'I got the idea that I should be covered by roses.' (Dagbladet 002)

- (e) *Jeg lager de filmene jeg har lyst å se selv.*
 I make the films+the I have desire to see myself
 'I make the films I want to see myself.' (Dagbladet 001)

Concerning the relationship between auxiliary and lexical verbs, we follow the principles described in 3.1.4.

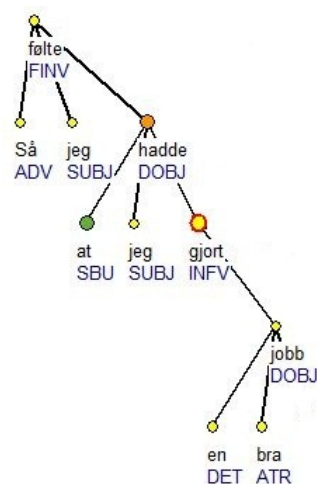
3.10.2 The complementizer

The complementizer must have the function SBU. Complementizers will always be dependent on the finite verb. In sentences like (3.119 a), consisting of a finite auxiliary verb and a non-finite lexical verb, the complementizer will, in other words, be dependent on the auxiliary, as the analysis in (3.119 b) shows:

- (3.119) (a) *Så jeg følte at jeg hadde gjort en bra jobb [...].*
 So I felt that I had done a good job

'So I felt that I had done a good job.' (Dagbladet 003)

(b)



Complementizers can sometimes be modified by an adverb, as in (3.120 a) and (3.120 b).

- (3.120) (a) *Selv om mange i Kr.F og Venstre ser ut til å*
 even if many in Kr.F and Venstre looks out to to
ha vanskelig for å innse det, er det et faktum at
 have difficult for to realize it is it a fact that

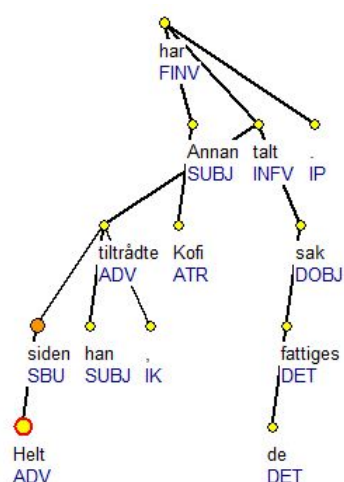
Fr.P. blir det viktigste støttepartiet for
 Fr.P. becomes the important.sup support party for
Bondeviks andre regjering.
 Bondevik's second government

'Even if many in Kr.F and Venstre seem to have difficulties realizing it, it is a fact that Fr.P. will become the most important support party for Bondevik's second government.' (Aftenposten 001)

- (b) *Helt siden han tiltrådte, har Kofi Annan talt de*
 even since he took office has Kofi Annan spoken the
fattiges sak.
 poor's cause
 'Even since he took office, Kofi Annan has adopted the cause of the poor.' (Aftenposten 001)

The adverb will be dependent on the complementizer with the function ADV, as (3.121) illustrates:

(3.121)



Quite often subordinate clauses occur without a complementizer. The complementizer *at* is often omitted in object clauses to some verbs, as in (3.122 a). Conditional clauses often occur with the verb in initial position and no complementizer, as in (3.122 b). Since the complementizer isn't the head, the analysis of (3.122 a) and (3.122 b) will be the same as for subordinate clauses with a complementizer.

- (3.122) (a) *Trur du Hellas vil greia dette?*
 believe you Greece will manage this
 'Do you believe Greece will manage this?' (Dag og Tid 002)

- (b) Held utviklinga fram, kjem Noreg til å gå
 hold development+the forward, comes Norway to to go
forbi Danmark i innbyggjartal i løpet av
 past Denmark in inhabitant number i course+the of
2020-talet.
 2020ies+the
 'If the development carries on, Norway will pass Denmark in number
 of inhabitants during the 2020ies.' (Dag og Tid 001)

3.10.3 Relative clauses

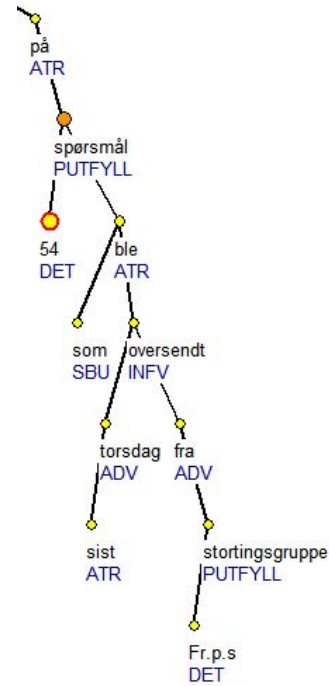
Relative clauses are clauses which modify a phrase in the superordinate clause, called the correlate. This correlate is also a part of the relative clause's clausal structure, where it is either explicitly expressed with a relative word (relative adverb, relative pronouns, etc.) or is implied. Relative clauses with an implied correlate often, but not always, have the complementizer *som*. This is illustrated by (3.123 a), (3.123 b) and (3.123 c), in which the correlate is emphasized and the relative clause is in brackets.

- (3.123) (a) *Der overleverte han de tre samarbeidspartienes*
 there delivered he the three coalition parties.gen+the
svar på 54 spørsmål [som sist torsdag ble
 answer to 54 questions that last Thursday became
oversendt fra Fr.p.s stortingsgruppe].
 sent over from Fr.p.'s parliamentary group
 'There, he handed over the three calition parties' answers to 54
 questions that were sent over to them from Fr.P.'s parliamentary
 group.' (Aftenposten 001)
- (b) *Han klipper opp klærne [han reiste i] [...].*
 he cut up clothes+the he traveled in
 'He cut up the clothes he traveled in' (Dagbladet 001)
- (c) *Hjelpemidlene er plater i rustfritt stål [der*
 remedies+the are plates in stainless steel where
budskapet er stanset ut med bokstaver].
 message+the is punched out with letters
 'The remedies are plates in stainless steel on which the message is
 punched out in letters.' (Dagbladet 001)

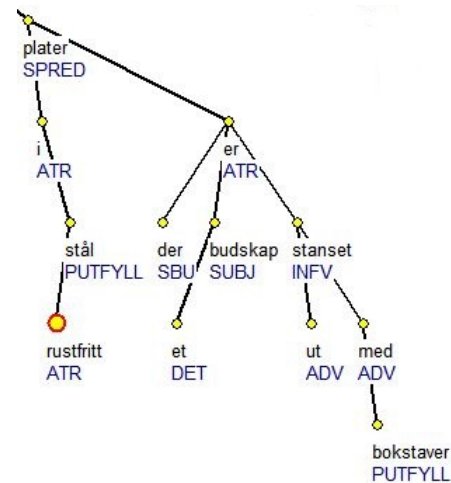
In (3.123 a) we have a relative clause with *som*. The correlate is the implied subject of the relative clause. In (3.123 b), the relative clause does not have a complementizer. The correlate is the implied complement to the preposition *i*. In (3.123 c), *der* is in the initial position of the relative clause. There are probably good linguistic reasons to count *der* and *da* at the beginning of relative clauses as relative words with a syntactic function in the subordinate clause. In the first part of the project period, we therefore annotated them as such. It turned out to give us some practical problems, and a syntactic parser we tested, had trouble finding a pattern in our analyzes. We therefore chose to analyze *der* and *da* in relative clauses as complementizers, and give them the same analysis as *som*.

We have chosen to let the finite verb in the relative clause be dependent on the correlate. If the correlate is a nominal phrase, the function will be ATR. *Som* will be dependent on the finite verb of the relative clause, like other complementizers are, with the function SBU. The analysis of (3.123 a) is therefore as in (3.124 a). The analysis of (3.123 c) is given in (3.124 b).

(3.124) (a)



(b)



The correlate of a relative clause is not necessarily a nominal phrase.¹⁷

¹⁷Faarlund et al. (1997, 1054-1057) admittedly restricts relative clauses to clauses with nominal correlate and use other terms for clauses with other types of correlates. Such a

The adverb *slik*, 'like that, such', often occurs with a relative clause, as in (3.125 a). Sometimes *slik* is implicit in the superordinate clause, as in (3.125 b). In the case of clefting, we also often have a non-nominal correlate. (3.125 c) is an example of clefting, where the prepositional phrase *i fylla* is the correlate. For more on clefting, see 3.3.

- (3.125) (a) - *Vi kunne ikke ha fortsatt slik [som det var].*
 we could not have continued such as it was

'We could not have continued the way it was.' (Dagbladet 003)

- (b) - *Dere gjør med vokalister [som eldre menn gjør med damer].*
 you do with vocalists as older men do with ladies

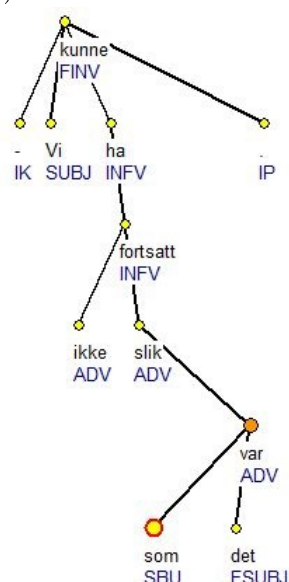
'You do to vocalists what older men do to ladies.' (Dagbladet 003)

- (c) - *Det er i fylla [jeg får høre det].*
 it is in drunkenness I get hear it

'It is when drinking I hear it.' (Dagbladet 001)

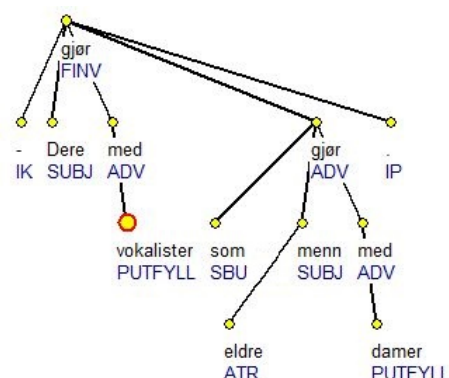
Modifiers of adverbs and prepositions should be ADV (cf. 3.5). This also applies to relative clauses. The relative clause in (3.125 a) shall be ADV on *slik*, as (3.126 a) shows. When the correlate is implicit in the superordinate clause, as in (3.125 b), the relative clause takes the correlate's function in the superordinate clause, cf. (3.126 b). In (3.125 c), the relative clause is dependent on the preposition with the function ADV, see (3.126 c).

- (3.126) (a)

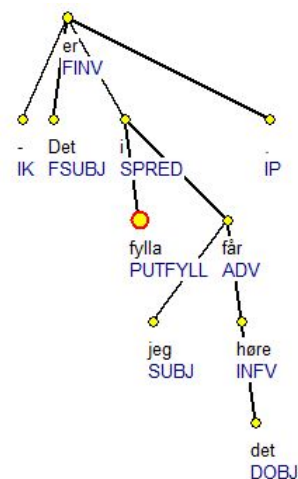


- (b)

distinction is not relevant in this context.



(c)



We have chosen not to distinguish between restrictive and non-restrictive relative clauses in the corpora. The non-restrictive relative clause in (3.127) should, in other words, have ATR, as other nominal relative clauses do:

- (3.127) Rigoberta Menchu, [*som fikk Nobels fredspris i 1992*],
 Rigoberta Menchu who got Nobel's peace price in 1992
takket nok en gang for det norske folks solidaritet
 thanked still a time for the Norwegian people's solidarity
med Guatemala.
 with Guatemala

'Rigoberta Menchu, who got Nobel's peace price in 1992, once again thanked for the solidarity of the Norwegian people with Guatemala.'
 (Aftenposten 001)

We also have non-finite relative clauses in Norwegian. More about this in section 3.7.

The complementizer *som* is used in more contexts than relative clauses. It is also used e.g. in comparative clauses, where no element in the superordinate clause is implied in the subordinate, cf. (3.128):

- (3.128) *Kjetil Hasselberg kjøper [...] nye vinterdekk samtidig*
 Kjetil Hasselberg buys new winter tires simultaneously
som han kjøper seg unna skiftinga.
 as he bought refl away changing+the
 'Kjetil Hasselberg buys new winter tires at the same time as he buys his way out of the changing of them.' (Dagbladet 004)

For more on such clauses, see 5.8.

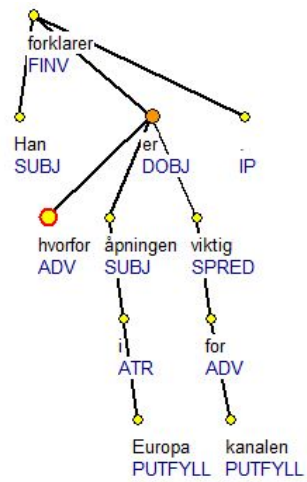
3.10.4 Indirect interrogative clauses etc.

In some contexts, a subordinate clause contain an interrogative word in its left periphery. The most common cases of this is indirect interrogative clauses, as in (3.129 a), but this also occurs in some other constructions, as in (3.129 b). Sometimes, especially when the interrogative word is a subject, the clause contains the complementizer *som*. The interrogative word then precedes the complementizer, as we see in (3.129 c) (for more on these constructions, see Faarlund et al., 1997, 989-994, 1052-1054, 1058-1061). The interrogative word is emphasized and the clause is put in square brackets in these examples.

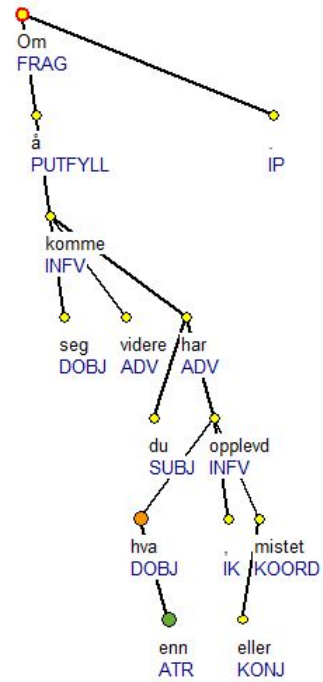
- (3.129) (a) *Han forklarer [hvorfor åpningen i Europa er viktig for kanalen].*
 He explains why opening+the in Europe is important for channel+the
 'He explains why the opening is important for the channel.'
 (Dagbladet 001)
- (b) *Om å komme seg videre [hva du enn har opplevd, eller mistet].*
 about to come refl wider what you ever have experienced or lost
 'About moving on, whatever you have experienced or lost.'
 (Dagbladet 001)
- (c) *En amerikansk journalist spør [hva som har vært høydepunktet i Stockholm].*
 an American journalist asks what that has been peak+the in Stockholm
 'An American journalists asks what the best moment has been in Stockholm.'
 (Dagbladet 001)

The analysis of (3.129 a) and (3.129 b) follows from what we have said earlier in this section: The finite verb is the head of the clause, the interrogative word is dependent on the word that it is an argument or adjunct of, which is not necessarily the finite verb. This is shown in (3.130 a) and (3.130 b). *Som* gets the regular analysis for complementizers, cf. (3.130 c):

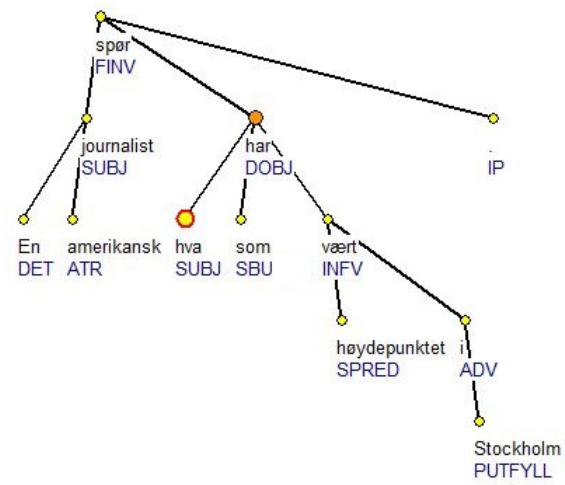
- (3.130) (a)



(b)



(c)



Chapter 4

Dependents with non-verbal heads

4.1 The nominal phrase

4.1.1 Introductory remarks

In this section we explain our analysis of nominal phrases (NPs)¹. A fundamental question here is what should be considered the head in complex nominal phrases - should it be the noun (or pronoun) or the determiner? There are good arguments for both solutions, but we cannot go into any further discussion of this here. In NDT, we have decided to let the noun be the head in cases where there is both a noun and determiner, in order to achieve the most consistent analyzes: Often nouns lack a determiner, and in such cases the noun would have to be the head anyway.

In NDT, we use three functions for dependent on nouns: ATR, DET and APP. ATR and DET will be described in more detail below in 4.1.2 and 4.1.3. The APP function has a greater application than just nominal phrases, and it has therefore received a special coverage in 4.2. In addition to describing the overall use of ATR, DET and APP, we mention in particular the analysis of numbers following a noun (see 4.1.4).

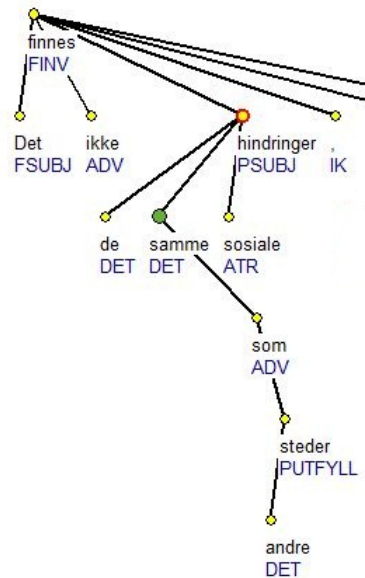
Before we go into details in the description of ATR and DET, we establish some general principles for the use of these two functions. ATR and DET equals roughly what Faarlund et al. (1997, 234) calls *beskrivere*, 'descriptors', and *bestemmere*, 'determiners'. The ATR function is used for *beskrivere*, typically adjectives, while the DET function is used for *bestemmere*, typically determiners. Most often, it will therefore be appropriate to analyze determiners as DET and adjectives as ATR when they are dependent on nouns. Exceptions to this will be explained in the following paragraphs. The example and diagram below illustrates the typical use of ATR and DET:

(4.1) (a) *Det finnes ikke de samme sosiale hindringer som andre*
it exists not the same social hindrance that other

¹As we described in 2.2, the phrase is not a fundamental concept in dependency grammar. When we still use the term *nominal phrase* here, it is to have a term that covers nouns/pronouns with their dependent.

steder [...]
 places
 'The same social hindrances are not found as other places.'
 (Klassekampen 001)

(b)



We see that the determiners *de* and *samme* is DET on the noun *hindringer*, while the adjective *sosiale* is ATR.

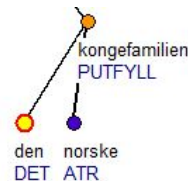
In many NPs containing both adjective and determiner plus noun, one can be in doubts about whether the determiner is dependent on the adjective or the noun. In the example below, should *de* be placed on *kongelig* or *bryllup*, and should *den* be placed on the *norske* or *kongefamilien*?

(4.2) (a) [...] *det kongelige bryllup*
 the royal wedding
 'the royal wedding' (Aftenposten 001)

(b) [...] *den norske kongefamilien* [...]
 the Norwegian royal family
 'the Norwegian royal family' (Aftenposten 001)

We have chosen a flat analysis, where both the adjective and the determiner is dependent directly on the noun, see the tree below:

(4.3) (a)



We will now look at some characteristics of ATR function.

4.1.2 ATR

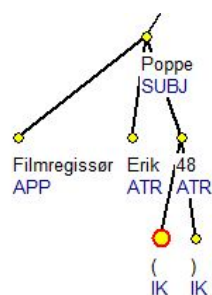
4.1.2.1 Different types of descriptors

ATR is used for descriptive dependents to nouns. We have seen that these often are adjectives. Another type of descriptors which often occurs in newspaper text, is age specifications and similar supplemental information, as in the examples below. These will also be counted as ATR (ATR-dependents are emphasized):

- (4.4) (a) *Filmregissør Erik Poppe (48) [...]*
 film director Erik Poppe (48)
 'Film director Erik Poppe (48)' (Dagbladet 001)
- (b) *Randi Sakshaug (Ap) [...]*
 Randi Sakshaug (Ap)
 'Randi Sakshaug (AP)' (Klassekampen 001)
- (c) *Poppe (t.v.) på jobb for VG [...]*
 Poppe (to the left) on work for VG
 'Poppe (to the left) working for VG' (Dagbladet 001)

The figure below shows the structure of (4.4 a):

(4.5)



Note that if the referent of the superordinate noun is referred to again, and not only assigned properties, we use APP and not ATR. The example below should not have ATR, but APP (see 4.2 for a detailed descriptions of the APP function):

- (4.6) *[...] tv-kanalen ESPN (Entertainment and Sports
 TV channel+the ESPN Entertainment and Sports
 Programming Network) [...]*
 Programming Network
 'the TV channel ESPN (Entertainment and Sports Programming
 Network)' (Dagbladet 001)

In addition to the dependents described above, prepositions are regularly dependents on nouns. The examples below illustrate this (NPs where nouns have preposition dependents, are emphasized):

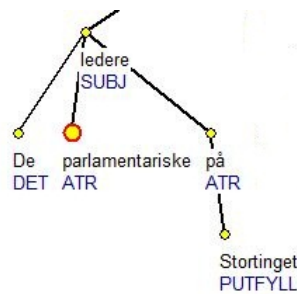
- (4.7) (a) *De parlamentariske ledere på Stortinget ble mandag
 the parliamentary leaders on Stortinget became Monday
 kveld enige (...)*
 evening agreed

'The parliamentary leaders in Stortinget agreed Monday evening.'
(Aftenposten 001)

- (b) [...] *stortingspresident Kirsti Kolle Grøndahl var*
President of Parliament Kirsti Kolle Grøndahl was
sikkert også fornøyd med storhertugen av Luxembourg
certainly also pleased with Grand duke+the of Luxembourg
'President of Parliament Kirsti Kolle Grøndahl was certainly pleased
with the Grans duke of Luxembourg' (Aftenposten 001)
- (c) *At Arbeiderpartiet får ledervervet i*
that Labour Party gets leader office in
utenrikskomitéen, er et signal om at
Foreign Affairs Committee+the is a signal about that
partiet forbereder seg på opposisjonstilværelsen.
party+the prepares refl on opposition existence+the
'That the Labour Party gets the office as leader og the Foreign
Affairs Committee, is a signal that the party prepares itself for being
in opposition.' (Aftenposten 001)

The figure below shows the structure of (4.7 a):

(4.8)



It can sometimes be difficult to determine whether a preposition should be ATR on a noun or ADV on a verb. An example of this can be seen here:

- (4.9) *Ap. får også lederverv i forsvarskomiteén og*
Ap. gets also leader offices in defence committee+the and
næringskomitéen.
industry committee+the
'The Labour Party also gets leader offices in the defence committee and
the industry committee.' (Aftenposten 001)

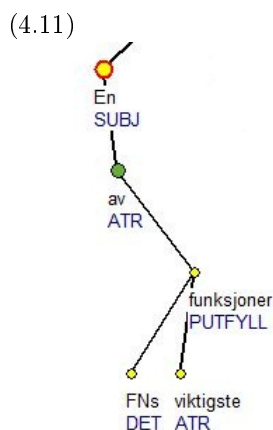
It is not obvious whether *i forsvarskommittéen og næringskommittéen* should be dependent on *får* or *lederverv*. In such cases of doubt, it may be helpful to try to move the noun and the preposition (with its complement) to the left periphery of the matrix clause. If this is possible without the sentence's meaning changing, it is an indication that the preposition should be dependent on the noun.

4.1.2.2 *Av-partitive*

A special use of ATR function is the construction we call the *av-partitive*: A determiner is followed by the preposition *av*, 'of', and its complement expressing a partitive relation.

- (4.10) (a) *En av FNs viktigste funksjoner [...]*
 one of UN's important.sup functions
 'One of the UN's most important functions' (Aftenposten 001)
- (b) *To av lagene fra verdens beste hockeyliga [...]*
 two of teams+the from world+the's best hockey league
 'Two of the teams from the world's best hockey league' (Dagbladet 001)

We let be the preposition be ATR on the the determiner:



As a rule, determiners take ADV-dependents (see 3.5). But we assume that, in the case of *av*-partitives, there is an NP with an elided noun, partly because of the obligatory concord in genus between the determiner and the prepositional complement. A nominal constituent such as *En* in the example above shall therefore be a determiner in the morphological annotation (not a pronoun). Because we assume that the preposition actually is dependent on an elided noun, it gets the function ATR.

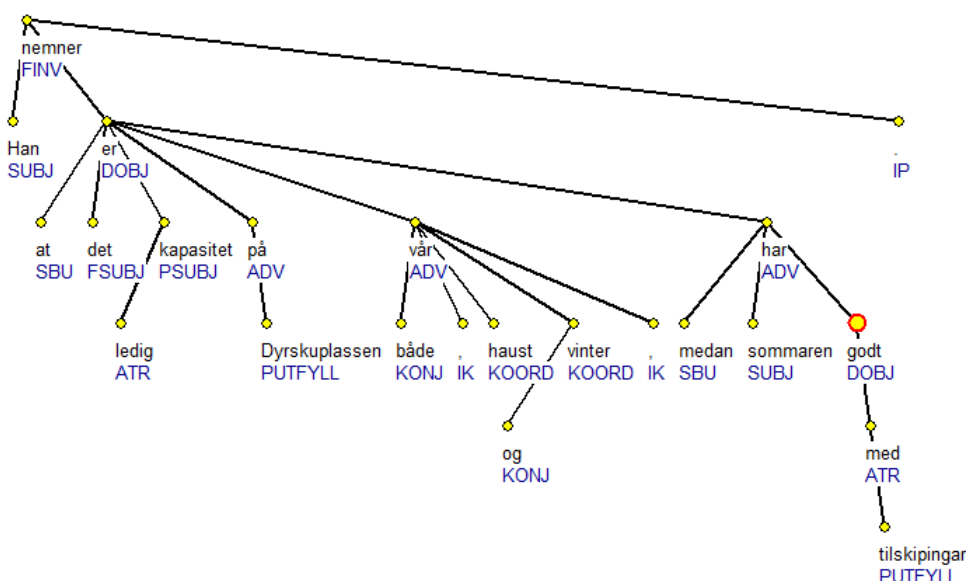
The analysis given above can be extended also to another partitive construction, namely construction with adjectives such as *rikelig*, 'abundant' and *godt*, 'good, here: a lot':

- (4.12) *Han fikk rikelig med fisk.*
 he got abundant with fish
 'He got a lot of fish.'

We will also assume here that we have an NP with an elided noun in these, and we therefore let the adjective *rikelig* serve as head. The prepositional phrase *med fisk* is considered as a complement to the elided noun. It becomes an ATR dependent on the adjective. An example of this analysis from the corpus is given below.

- (4.13) (a) *Han nemner at det er ledig kapasitet på
 he mentions that it is available capacity on
 Dyrsakuplassen både vår, haust og vinter, medan
 Dyrsakuplassen both spring fall and winter while
 sommaren har godt med tilskipingar.
 summer+the has good with events
 'He mentions that there is available capacity at Dyrsakuplassen
 spring, fall and winter, while the summer has a lot of events.'
 (Vest-Telemark Blad 004)*

(b)



Finally, it should be mentioned that a common dependent on nouns/pronouns is relative clauses. Here, we also use ATR - see 3.10.3 for more on finite relative clauses and 3.7.2 for more on infinitival relative clauses. We also use the ATR for participles following a noun - see 3.7.4 for more on this.

4.1.3 DET

As mentioned earlier, determiners usually have the function DET. DET is used, however, in a few other contexts as well, which we will describe in more detail in the following paragraphs. In 4.1.4, we mention an exception to the rule that determiners should be DET.

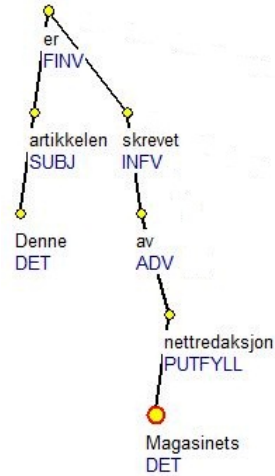
4.1.3.1 Genitive

Genitive nouns should be dependent on the noun they modify, with the function DET, cf. sentence (4.14 a) and the analysis in (4.14 b):

- (4.14) (a) *Denne artikkelen er skrevet av Magasinets
 this article+the is written of Magasinets's
 nettredaksjon [...].
 web editorial office*

'This article is written by Magasinet's web journalists.' (Dagbladet 001)

(b)



Note that the genitive *s* is attached to the last word in the phrase constitutes the genitive expression, and that this word is not necessarily the head. In (4.15 a), the last word in a complex proper noun has the genitive *s*. The first word, however, is the head, according to our analysis (cf. 4.5.3). It is therefore the first name that will get the function DET, as shown in (4.15 b).

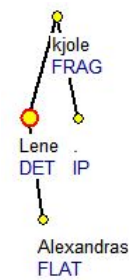
(4.15) (a)

Lene Alexandras *kjole*.

Lene Alexandra's dress

'Lene Alexandra's dress.' (Dagbladet 002)

(b)



4.1.3.2 Quantity nouns

In Norwegian, a noun indicating a quantity, such as *del* in (4.16 a), *rekke* in (4.16 b) and *injeksjon* in (4.16 c), can immediately precede a noun:

(4.16) (a) *Det er en del ting du bør huske på [...].*
 it is a part things you should remember on

'There are a few things you should remember' (Dagbladet 004)

- (b) *Et kamera vi filme en rekke prostituerte ved Halmtorvet*
 a camera will film a row prostitutes by Halmtorvet
 [...].

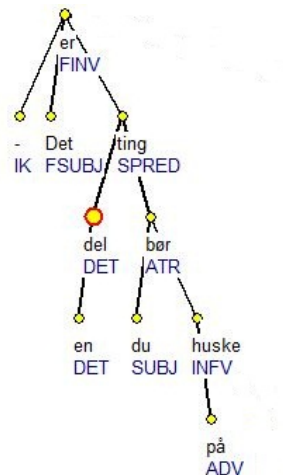
'A camera will film several prostitutes at Halmtorvet.' (Dagbladet 004)

- (c) *Av og til blir det så ille at jeg trenger en*
 of and to becomes it so bad that I need an
 injeksjon kortison [...].
 injection cortisone

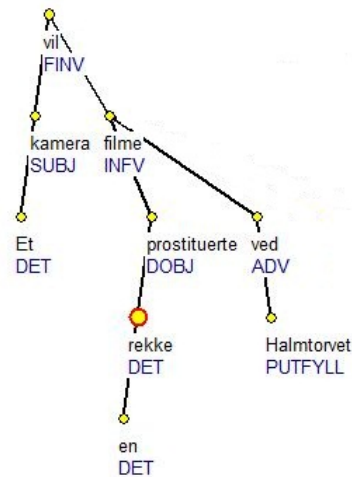
'Sometimes, it becomes so bad that I need an injection of cortisone.' (Dagbladet 003)

It is not obvious which of the two nouns should be the head. We have chosen to follow Faarlund et al. 1997, 236-240, who considers the quantity noun as a kind of quantifier to the following noun. In our analysis, the quantity noun gets the function DET. The analysis of (4.16 a), (4.16 b) and (4.16 c) is therefore as follows:

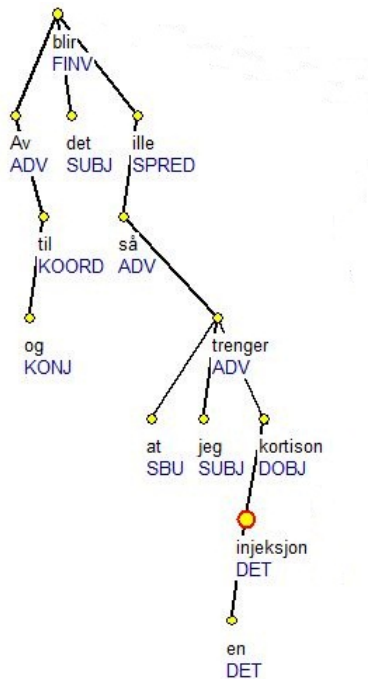
(4.17) (a)



(b)



(c)

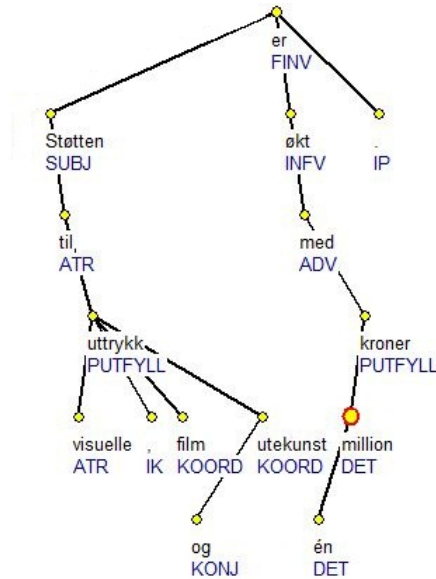


Number nouns such as *hundre*, 'hundred', *tusen*, 'thousand' *millioner*, 'millions' etc. are analyzed the same way:

- (4.18) (a) [...] *Støtten til visuelle uttrykk, film og*
 support+the to visual expressions film and
utekunst er økt med én million kroner.
 outdoor art is increased with one million kroner

'The support for visual expressions, film and outdoor art has been increased by one million kroner.' (Klassekampen 001)

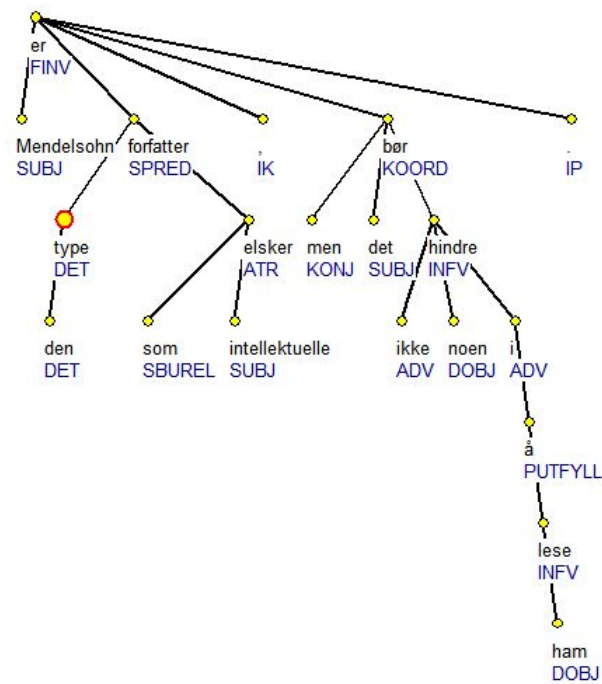
(b)



Nouns such as *type*, 'type', *slag*, 'type' etc., in examples such as (4.19 a) seem to have a lot in common with quantity nouns. The analysis is therefore the same, cf. (4.19 b):

- (4.19) (a) *Mendelsohn er den type forfatter som intellektuelle elsker,*
 Mendelsohn is the type author who intellectuals love
men det bør ikke hindre noen i å lese ham.
 but it should not prevent anyone in to read him
 'Mendelsohn is the type of author who intellectuals love, but that
 should not prevent anyone from reading him.' (Aftenposten 012)

(b)

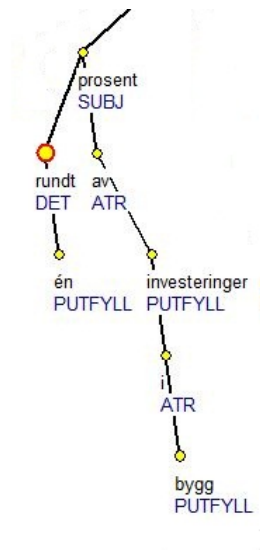


4.1.3.3 Quantity designations with a preposition

The prepositions *over* , 'over', and *rundt* , 'around', are often used in quantity designations, such that in (4.20 a). In such cases, the preposition should be DET on the noun and take the determiner as complement, as shown in (4.20 b):

- (4.20) (a) [...] *rundt én prosent av investeringer i bygg skal brukes på utsmykning [...].*
 around one percent of investments in buildings shall
 use.pass on decoration
 'Around one percent of investments in building must be used on
 decoration' (Klassekampen 001)

(b)



When such quantity designations with preposition co-occurs with a quantity noun, it is still only the determiner that will be complement to the preposition:

- (4.21) (a) *Det er bra, men ikke bra nok for et land med*
 it is good but not good enough for a country with
over 26 millioner innbyggere.
 over 26 million inhabitants
 'This is good, but not good enough for a country with over 26
 million inhabitants.' (Dagbladet 004)

(b)



See also 4.3.1.

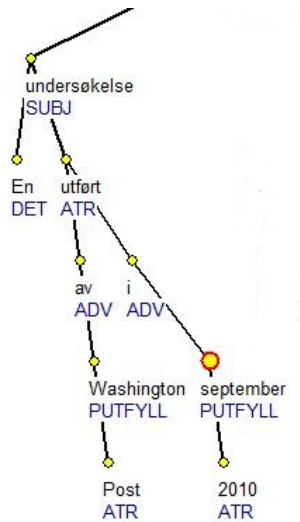
4.1.4 Cardinal numbers following a noun

Cardinal numbers are determiners, and they will therefore usually be DET-dependents. In the case where a number indicates e.g. a year, as in (4.22), we don't think the standard analysis should apply:

- (4.22) *En undersøkelse utført av Washington Post i september 2010 [...].*
 a survey accomplished of Washington Post in
 september 2010 [...].
 September 2010
 'A survey carried out by the Washington Post in September 2010'
 (Klassekampen 001)

Cardinal numbers in such cases have no quantifying function, unlike cardinal numbers preceding the noun: *2010* in (4.22) does not indicate a quantity. We believe, based on this, that there is reason not to count them as determiners (cf. Faarlund et al. 1997, 233-234). Therefore, these should be ATR-dependents on the preceding noun:

(4.23)

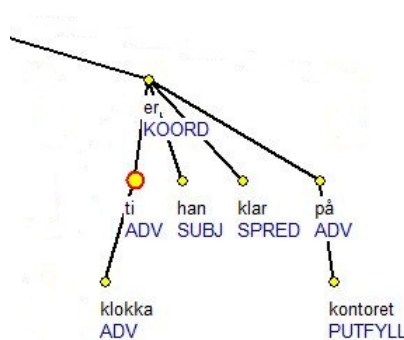


Klokka, 'clock+the', and *nummer*, 'number', are also often followed by cardinal numbers, as in (4.24 a) og (4.24 b).

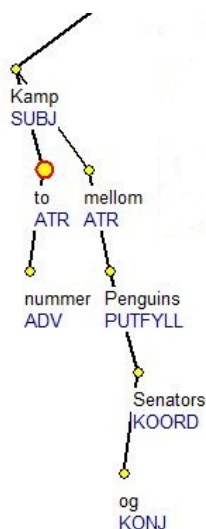
- (4.24) (a) *[...] klokka 10 er han klar på kontoret.*
 clock+the 10 is he ready on office+the
 'At 10 o'clock, he is ready in the office.' (Dagbladet 003)
- (b) *Kamp nummer to mellom Penguins og Senators [...].*
 match number two between Penguins and Sentators
 'Match number two between Penguins and Sentators' (Dagbladet 001)

These cases differ from (4.22), because *klokka* and *nummer* can be removed without the sentence being ungrammatical or the meaning changing, while the cardinal numbers cannot be removed. Therefore, we believe it is appropriate to make *klokka* and *nummer* dependents on the numbers. They get the function ADV:

(4.25) (a)



(b)



4.2 APP

This section is a discussion of the use of the APP function. APP stands for *apposition*, which can be described as a constituent that has the same reference as another constituent in the sentence. Faarlund et al. (1997, 270-274) reserves the term apposition of nominal constituent, but we will use it in a broader sense.

4.2.1 below is a discussion of APP as nominal apposition, i.e. apposition in the traditional sense. In NDT, we have separate rules for the analysis of proper names preceded job titles and the like. This is described in 4.2.2. We will further explain how we use APP when an element is duplicated in the sentence (4.2.3), in connection with the adverb *så* (4.2.4), and in a construction with *noe* and a relative clause (4.2.5).²

²In the latter case, APP is usually a dependent on the verb. Although the current chapter is titled *Dependents with non-verbal heads*, it seems natural to mention the APP-dependents in 4.2.5 here.

In 4.2.6, we discuss appositions with the abbreviation *dvs.*, while 4.2.7 describes the analysis of appositions that specify a subset.

4.2.1 APP in nominal phrases

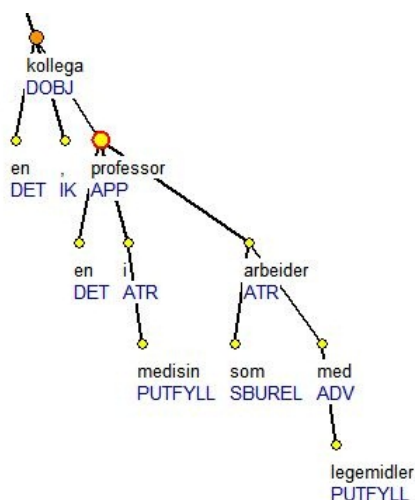
The analyzes presented here applies to appositions with common nouns. We have a separate analysis of apposition with proper nouns, which will be presented in 4.2.2.

We use the APP function in cases where two (or more) phrases with the same reference together form a larger nominal phrase, as in the emphasized phrases below:

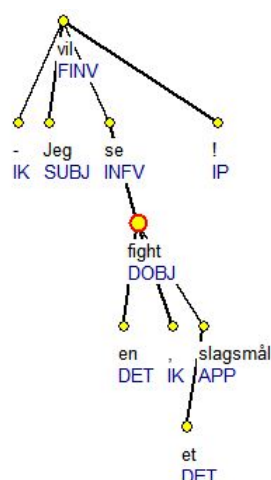
- (4.26) (a) *Jeg stusset og ringte en kollega, en professor i
I wondered and called a colleague a professor in
medisin som arbeider med legemidler.
medicine who works with drugs
'I wondered and called a colleague, a professor in medicine who
works with drugs.'* (Aftenposten 011)
- (b) *Jeg vil se en fight, et slagsmål!
I will see a fight a battle
'I want to see a fight, a battle!'* (Dagbladet 001)

One of the two nominal constituents is head, while the other one becomes an APP-dependent on it. It can be difficult to determine which is the head in such constructions, but for the sake of uniqueness, we let the first constituent (in the linear structure) be the head. In example (4.26 a), *kollega* is therefore the head and *professor* is APP, while *fight* is head and *slagsmål* dependent in (4.26 b):

- (4.27) (a)



- (b)



Nominal constituents can often be further specified by an apposition. The analysis we have presented above, also applies for cases like these:

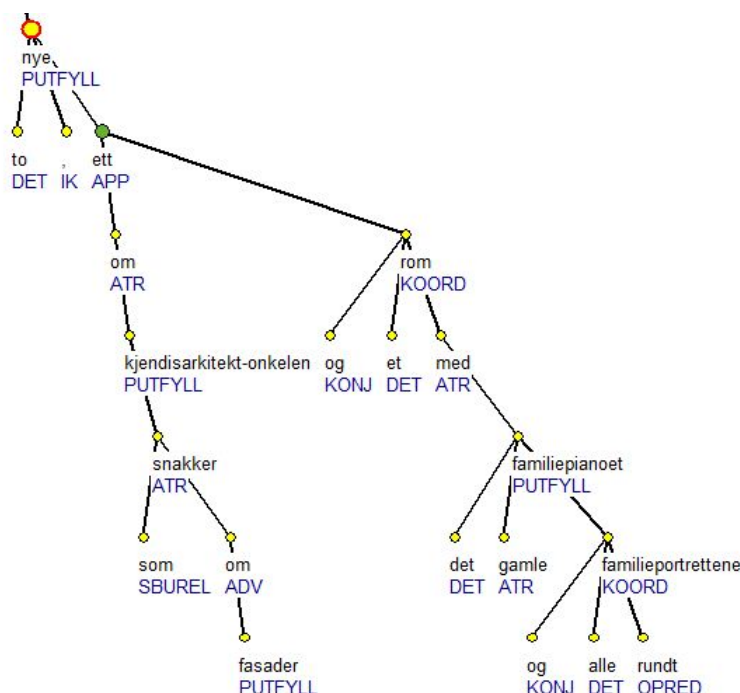
- (4.28) (a) *Gjennom 20- og 30-tallet oppsto dermed de*
 through 20ies and 30ies arised consequently the
såkalte standarddansene, vals, tango, wienervals, slowfox
 so-called standard dances wals tango Viennese wals slowfox
og quickstep, *i sine moderne former.*
 and quickstep in their modern forms
 'During the 20ies and 30ies, the so-called standard dances came into
 existence, wals, tango, Viennese wals, slowfox and quickstep, in their
 modern forms.' (Dagbladet 001)

(b) ()

I tillegg til to nye, ett om
 in addition to two new one about
kjendisarkitekt-onkelen som snakker om fasader og
 famous architect uncle+the who talks about fasades and
et rom med det gamle familiepianoet og alle
 a room with the old family piano+the and all
familieportrettene rundt.
 family portraits+the around
 'In addition to two new ones, one about her famous architect uncle
 who talks about fasades, and about a room with the old family piano
 and with alle the family portraits around it.' (Dagbladet 003)

The structure of the relevant part of (4.28 b) is reproduced below. Note that the APP-phrase consists of two conjuncts. (See 4.4 for more on coordination.)

(4.29)



A clause with the complementizer *at* or an infinitive following the pronoun *det*, will also be analyzed as appositions.

- (4.30) - *Det gjelder både kostymene og det at de løfter*
 it applies both costumes+the and that that they lift
hverandre [...]
 each other
 'This applies both to the costumes and to the fact that they lift each other.' (Dagbladet 002)

In such constructions, *det* becomes the head and the *at*-clause an APP-dependent.

4.2.2 Special rules for APP with proper nouns

We had a bit of trouble finding a suited analysis for cases where a proper name has a (common) noun in front of it and the two are in an apposition relation to each other. We could have analyzed appositions with common and proper nouns alike, but then we run into difficulties with titles to person names: In cases like *kong Harald*, 'King Harald', *forfatter Erling Loe*, 'author Erling Loe', and *filmregissør Erik Poppe*, it is not intuitive to let the title be the head and the person name be a dependent. In the case of proper names that are not person names, as in the following cases, a standard apposition analysis would have been adequate, however:

- (4.31) (a) *I slutfasen av valgkampen toppet*
 in end phase+the of election campaign+the topped
partiet Høyre optimismen rundt hvordan vi kan få
 party+the Høyre optimism+the around how we can get

større innslag av blandingsøkonomi i norsk kunstliv
 bigger element of mixed economy in Norwegian art life
 [...]

'In the end phase of the election campaign, the party Høyre topped the optimism concerning how we can get a greater element of mixed economy in Norwegian art life.' (Aftenposten 012)

- (b) *Erindringen inntar også en sentral plass i*
 recollection+the takes also a central space in
 Mendelsohns debutbok: The Elusive Embrace [...]
 Mendelsohn's debut book The Elusive Embrace

'Recollection is also an important aspect of Mendelsohn's first book: The Elusive Embrace' (Aftenposten 012)

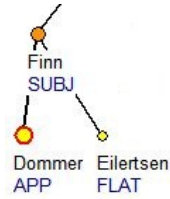
In an early phase of the project, we had a separate analysis of person names. This had to change, however, as it is problematic for a number of users of the corpus that we distinguish between different types of proper nouns. It was difficult to find other appropriate criteria. One could imagine that the title analysis should only be used if the first noun was indefinite, but this does not work, partly because of examples like (4.31 b). Instead of finding criteria other than the distinction between personal names and other proper nouns, we decided to use a unified analysis for all cases of a noun and a proper noun which are in an apposition relation to each other and accept that this would have some unintended consequences. The special analysis of proper names is as follows:

1. If a nominal phrase (with a common noun as head) immediately precedes a proper noun, the proper noun will be the head.
2. If the two nominal phrases are not immediately next to each other, the standard apposition analysis should be followed, cf. 4.2.1 and 4.2.3.
3. If the first nominal element is not a noun, but e.g. a pronoun, we follow the standard apposition analysis. The standard analysis is also followed when a proper noun is followed by another.

As a consequence of point 1, the analysis of the emphasized phrases in (4.32 a) will be as shown in (4.32 b):

- (4.32) (a) *Dommer Finn Eilertsen avstår, selvfølgelig bevisst,*
 judge Finn Eilertsen refrains of course consciously
fra å «sette ord på» det inntrykk retten for
 from to put words on the impression court+the for
sitt vedkommende måtte ha dannet seg av de handlinger
 its part would have formed refl of the actions
retten finner bevist og av lovovertrederen.
 court+the finds proven and of offender+the
 'Judge Finn Eilertsen abstains, consciously of course, from
 formulating the impression the court, for their part, would have
 formed of the actions the court finds have been proven, and of the
 offender.' (Aftenposten 012)

(b)

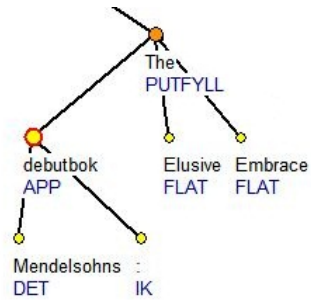


(4.31 a) and (4.31 b) get the following analysis, as a consequence of point 1:

(4.33) (a)



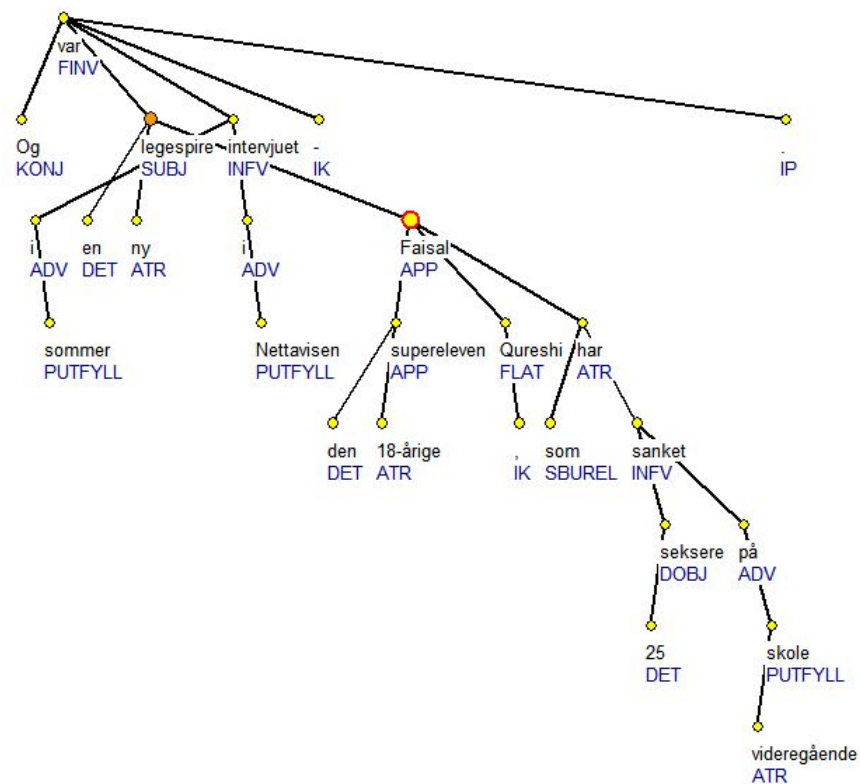
(b)



In (4.34 a), there are several words between the two nouns which are in an apposition relation to each other. In accordance with point 2, the standard apposition analysis applies, i.e. the first noun is the head:

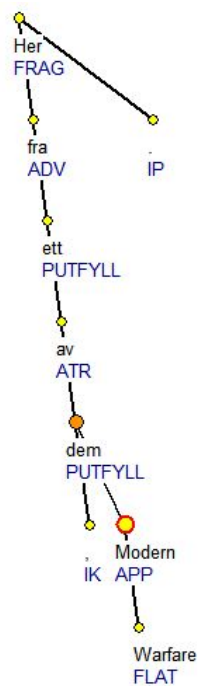
- (4.34) (a) *Og i sommer var en ny legespire intervjuet i Nettavisen - den 18-årige supereleven Faisal Qureshi, som har sanket 25 seksere på videregående skole.*
 and in summer was a new doctor sprout interviewed in Nettavisen the 18 year old super-pupil Faisal Qureshi who has collected 25 six-s in higher school
 'And this summer, a new prospective doctor was interviewed in Nettavisen - the 18 year old super-pupil Faisal Qureshi, who has collected 25 6-grades in high school.' (Aftenposten 009)

(b)



In (4.35 a), there is an apposition relation between a pronoun and a proper name. According to point 3, the standard apposition analysis applies - the pronoun becomes the head:

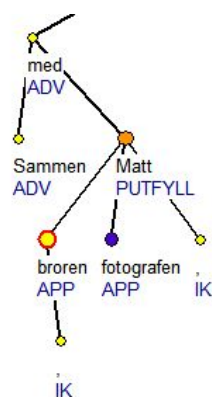
- (4.35) (a) *Her fra ett av dem, Modern Warfare.*
 here from one of them Modern Warfare
 'Here from one of them, Modern Warfare.' (Klassekampen 005)
- (b)



If there are several appositions to a proper noun, they all become dependents on the proper noun:

- (4.36) (a) *Sammen med broren, fotografen Matt, reiste*
 together with brother+the photographer+the Matt traveled
han i fem år rundt på tre kontinenter [...]
 he in five years around on three continents
 'Together with his brother, the photographer Matt, he traveled for
 five years around on five continents.' (Aftenposten 012)

(b)



4.2.3 APP when dislocated positions are duplicated

Sometimes, we encounter sentences where a phrase is represented twice, once in its normal position and once in what we will consider an adjoint (dislocated)

position at the beginning or the end of the sentence. In the examples below, the duplicated phrases are emphasized:

- (4.37) (a) - At andre regissører kaller deg pompøs; *kan* det *også*
 that other directors call you pompous can that also
 skyldes misunnelse?
 cause.pass envy
 'That other directors call you pompous; can that also be due to
 envy?' (Dagbladet 001)
- (b) Du *er flink*, du!
 you are good you
 'You, you are good!'
- (c) Da *var det fint*, da.
 then was it nice then
 'Then, then it was nice.'
- (d) Denne boka, den *bør du lese*.
 this book it should you read
 'This book, you should read it.'

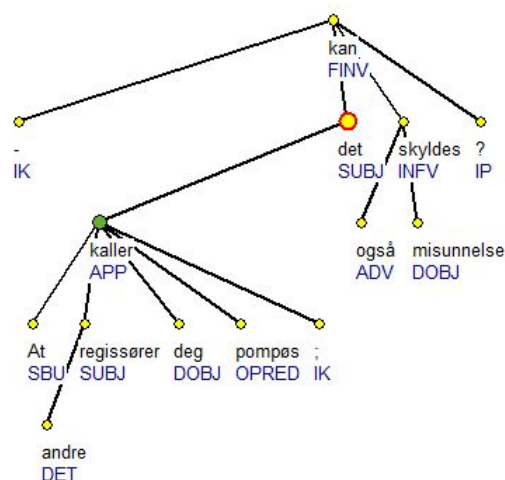
In (4.37 a), the clause *At andre regissører kaller deg pompøs* is left-dislocated and taken up again by the pronoun *det*. In (4.37 b), the subject *du* is repeated at the end of the sentence. This kind of duplication occurs primarily with subjects, but as we see in (4.37 c), it is also possible with some adverbs, such as *da* (the analysis assumes that the last *da* actually has a temporal content, as the first *da*, and not simply a discourse particle). In (4.37 d), we have *denne boka* both in a left-dislocated position at the very beginning of the sentence and in the shape of the pronoun *den* in the position directly before the finite verb.

Argument functions can never be represented more than once in a sentence, and neither can adjunct functions when they have the same referent. Therefore, we allow only one of the two phrases to be dependent on the sentence's head, with the respective syntactic function (for example SUBJ, as in (4.37 a), or DOBJ, as in (4.37 d)). The remainder of the two phrases are APP dependent on the former.

We let the phrase that is in its normal position in the sentence be head. The second phrase, which is dislocated at the beginning or end of the sentence, get the APP function.

Below we see (4.37 a) repeated with full analysis:

- (4.38) (a) - At andre regissører kaller deg pompøs; *kan* det *også*
 that other directors call you pompous can that also
 skyldes misunnelse?
 cause.pass envy
 'That other directors call you pompous; can that also be due to
 envy?' (Dagbladet 001)
- (b)



4.2.4 APP with *så*

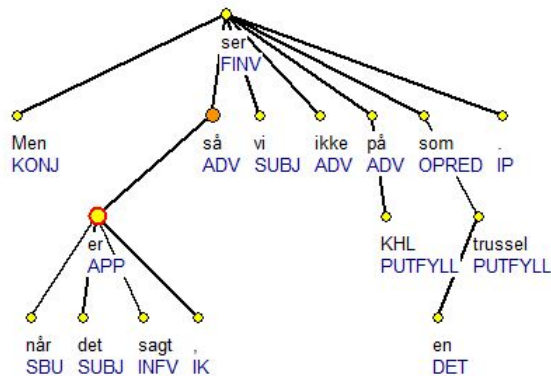
It is not uncommon that an adverbial phrase in the left periphery of the sentence is followed by the adverb *så*, which in turn precedes the finite verb.

- (4.39) (a) *Men når det er sagt, så ser vi ikke på KHL som en trussel.*
 bot when that is said so see we not on KHL as
 a threat
 'But, when this is said, we don't see KHL as a threat.' (Dagbladet 001)
- (b) *Om vi bruker store ord om hverandre nå, så er det fordi vi vet å sette pris på hverandre.*
 if we use big words about each other now so is it because we know to put price on each other
 'If we use big words about each other now, it is because we truly value each other.' (Dagbladet 001)
- (c) *Som du sikkert vet, så kommer over 30 prosent av spillerne i NHL fra Europa.*
 as you surely know so comes over 30 percent of players+the in NHL from Europe
 " (Dagbladet 001)

In such cases, we expect that *så* is co-referential with the preceding, adverbial phrase - it is otherwise difficult to see what *så* shall refer to. (We are talking here of course not about sentences where *så* has the meaning 'then'.) Since *så* repeats the contents of the preceding phrase, we use an apposition analysis in line with what we have described for duplicated subjects above: *Så*, which is in the position before the finite verb, i.e. a common position for adverbial phrases, gets the function ADV, while the preceding phrase becomes dependent on *så* with the function APP.

The figure below shows the structure of (4.39 a):

(4.40)



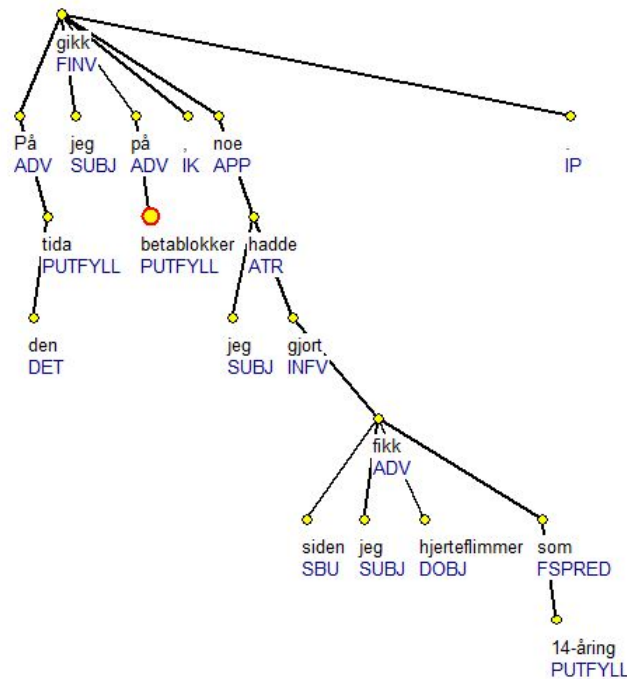
4.2.5 APP with *noe*

Sometimes, the content of a sentence is repeated by the pronoun *noe*, 'something', which is modified by a relative clause:

- (4.41) *På den tida gikk jeg på betablokker, noe jeg
 on that time+the went I on beta blocker something I
 hadde gjort siden jeg fikk hjerteflimmer som 14-åring.
 had done since I got heart flicker as 14-year-old
 'At that time, I used beta blockers, something which I had done since I
 got heart flicker when I was 14.'* (Dagbladet 001)

We let *noe* be an APP-dependent on the head of the clause to which it refers, whereas relative clause is an ATR-dependent on *noe*. The figure below shows the analysis of (4.41):

(4.42)



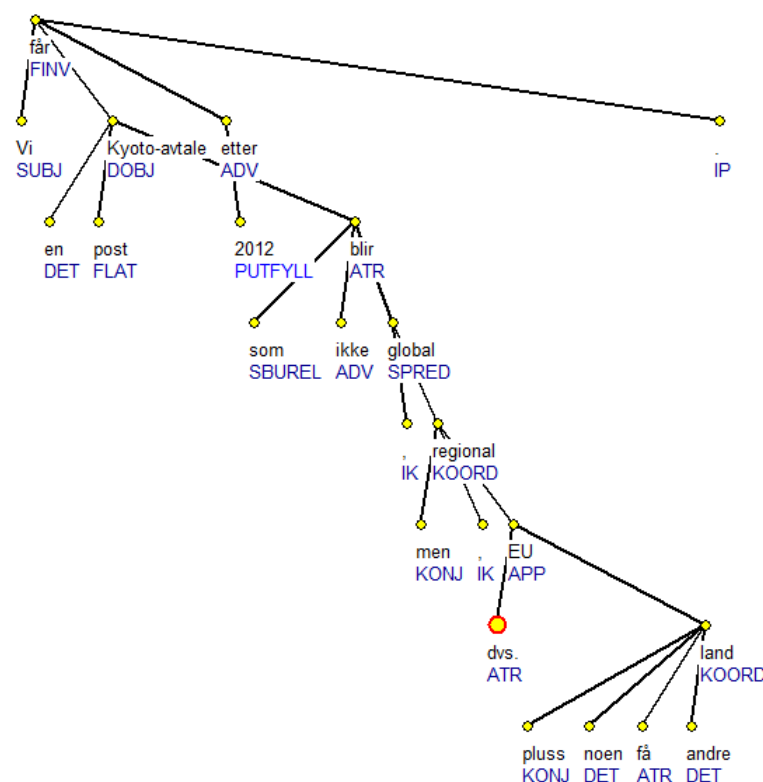
4.2.6 Appositions with the abbreviation *dvs.*

When an apposition is preceded by the abbreviation *dvs.*, 'i.e.', we disregard the words that the abbreviation stands for, and analyzes it as an adverb - i.e. it is dependent on apposition, and gets the functions ADV or ATR

- (4.43) (a) *Vi får en post Kyoto-avtale etter 2012 som ikke
 we get a post Kyoto treaty after 2012 which not
 blir global, men regional, dvs. EU pluss noen få
 becomes global but regional i.e. EU plus some few
 andre land.
 other countries*

'We will get a post-Kyoto treaty after 2012 which will not be global, but regional, i.e. the EU and a few other countries.' (Aftenposten 004)

(b)



4.2.7 Subset appositions

In NRG, a paragraph on appositions mentions briefly that [...] the *apposition* can also be an adjunct which elaborates on or clarifies the content of the head. [...] (Faarlund et al. (1997, 271)):

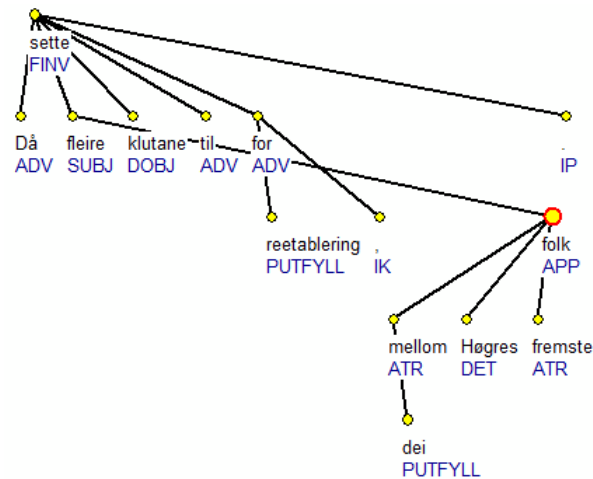
- (4.44) *ein delegasjon på fem personar, blant dei*
 a delegation on five persons among them
statsministeren
 prime minister+the
 'a delegation of five people, among them the prime minister' Faarlund
 et al. (1997, s. 271)

In our work with the corpus, we have discovered that the appositions of this type, indicating subsets (e.g. *statsministeren*) of the head word (e.g. *fem personar*), is a more complex subject, not least syntactically speaking, than what the reference grammar indicates. Here, we will refer to such appositions as subset appositions, though, as will become clear below, there are cases where such apposition refers to the entire core they are dependent on.

The simplest types subset appositions are those preceded by adverbial phrases of the type *blant annet*, 'including', *blant andre*, 'among others', *blant dei*, 'among them', *mellom anna*, 'including' etc. In these constructions, the adverbial phrase attaches an ADV-dependent on the apposition (and if the apposition is an NP, which will often be the case, the function ADV will be replaced with ATR, according to the standard practice for modifiers).

- (4.45) (a) *Då sette fleire klutane til for reetablering, mellom dei Høgres fremste folk.*
 then put more rags on for reestablishment, among
 them Høyre's leading people
 'Then, more people made an effort for reestablishing it, among them
 Høyre's leading people.' (Vest-Telemark Blad 005)

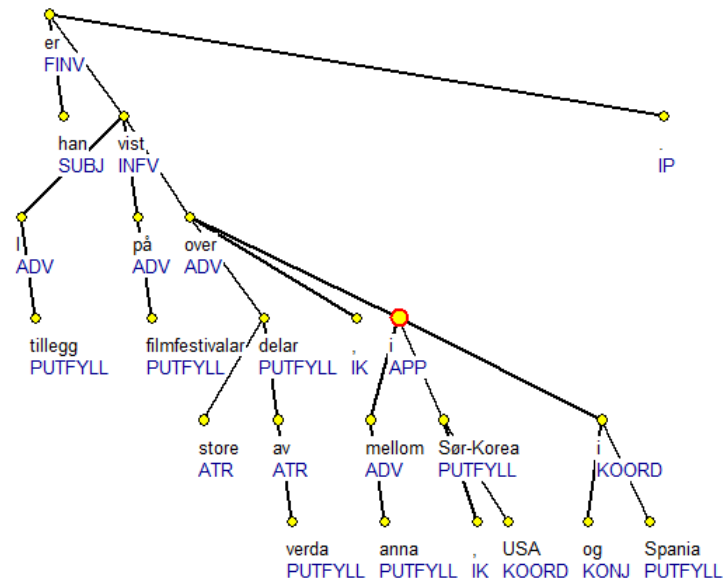
(b)



- (c) *I tillegg er han vist på filmfestivalar over store delar av verda, mellom anna i Sør-Korea, USA og i Spania.*
 in addition is he shown on film festivals over big parts
 of world+the between others in South Korea USA and in
 Spain

'It is also shown on film festivals many places worldwide, including in
 South Korea, USA and Spain.' (Dag og Tid 007)

(d)



What is not mentioned in NRG, is that the element before the apposition does not need to be an adverbial phrase, but may also be a non-adverbial NP:

- (4.46) *Ingen ble skadet da politiet pågrep 75 personer,*
 no one was hurt when police+the arrested 75 persons
tre av dem nordmenn, for å hindre bråk.
 three of them Norwegians for to hinder trouble
 'No one was hurt when police arrested 75 people, three of them
 Norwegians, to prevent trouble.' (p4.no, Google)

- (4.47) *To svenske pastorer, den ene av dem en kvinne, må*
 two Swedish pastors the one of them a woman must
møte i retten i Sverige.
 meet in court+the in Sweden
 'Two Swedish pastors, one of them a woman, must meet in cort in
 Sweden.' (Webavisen, Google)

In such cases, the apposition has the form of a small clause, i.e. a syntactic structure consisting of a subject (*tre av dem*; *den ene av dem*) and a non-verbal predicate (*nordmenn*, *en kvinne*) but without any finite verb. This is reminiscent of the structures we get with coordination and verb ellipsis (5.1.3), but we still choose not to use KOORD-ELL in the analysis here, because (1) these structures may, as in examples above, occur within another clause, clearly linked syntactically to the preceding NP, causing them to appear more like appositions than as independent clauses, and (2) even though they may be in sentence-final position (*To svenske pastorer må møte i retten i Sverige, den ene av dem en kvinne*), the sentence becomes ungrammatical if you add a conjunction (** To svenske pastorer må møte i retten i Sverige, og den ene av dem en kvinne*).

We will therefore also here use a subset apposition analysis. The NPs that provides detailed information about the subset will be the apposition (as *statsministeren* and *Høgres fremste folk* has same semantic function in example (4.44) and (4.45 a)): *nordmenn* and *kvinne* are heads for the appositions in

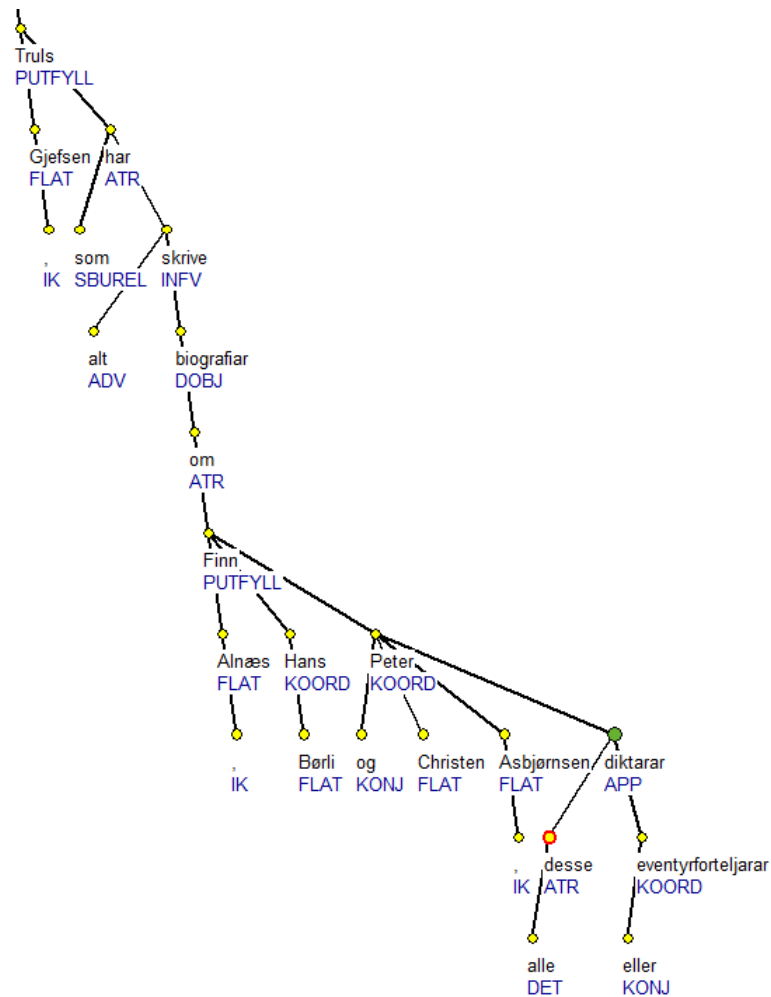
(4.46) and (4.47), and gets the function APP and are dependents on *personer* and *pastorer*, respectively.

But what do we do with the NPs which delimit the subset about which the apposition provides information, i.e. *tre av dem* and *den ene av dem* in (4.46) and (4.47)? It is not obvious what the answer is, as these NPs, as explained above, functions as subjects, regardless of finite verbs. Presumably there are a number of possible syntactic analyzes of this, but here we will go outside of this linguistic minefield and instead pursue a pragmatic solution that is most parallel with the analysis we have given to other appositions. In the examples (4.44) and (4.45 a), it is the adverbial phrases *mellom anna* and *mellom dei* which delimit the subset, and these are ADV/ATR-dependents on the apposition. We will therefore do the same with NPs like *tre av dem* and *den ene av dem* - we will analyze them as delimiting adverbial phrases without a preposition. Basically, they will have the function ADV, or ATR when they are dependent on NPs - as in the cases above, where they would be ATR-dependents on *nordmenn* and *kvinne* respectively.

In the corpus, there are just a few examples of subset appositions of this kind - and some of these examples show that the apposition does not necessarily need to indicate an actual subset, but rather the entire set. In (4.48 a), we find the subset apposition *alle desse diktarar eller eventyrforteljarar*. The delimiting NP - *alle desse* - is used to explicitly say that the apposition counts for the entire set that the apposition is dependent on (*Finn Alnæs, Hans Børli og Peter Christen Asbjørnsen*).

- (4.48) (a) *Truls Gjefsen, som alt har skrive biografier om*
 Truls Gjefsen who already has written biographies about
 Finn Alnæs, Hans Børli og Peter Christen Asbjørnsen,
 Finn Alnæs Hans Børli and Peter Christen Asbjørnsen
 alle desse diktarar eller eventyrforteljarar.
 all these poets or fairy tale tellers
 ‘Truls Gjefsen, who already has written biographies about Finn
 Alnæs, Hans Børli and Peter Christen Asbjørnsen, all of them poets
 or fairy tale tellers.’ (Dag og tid 004)

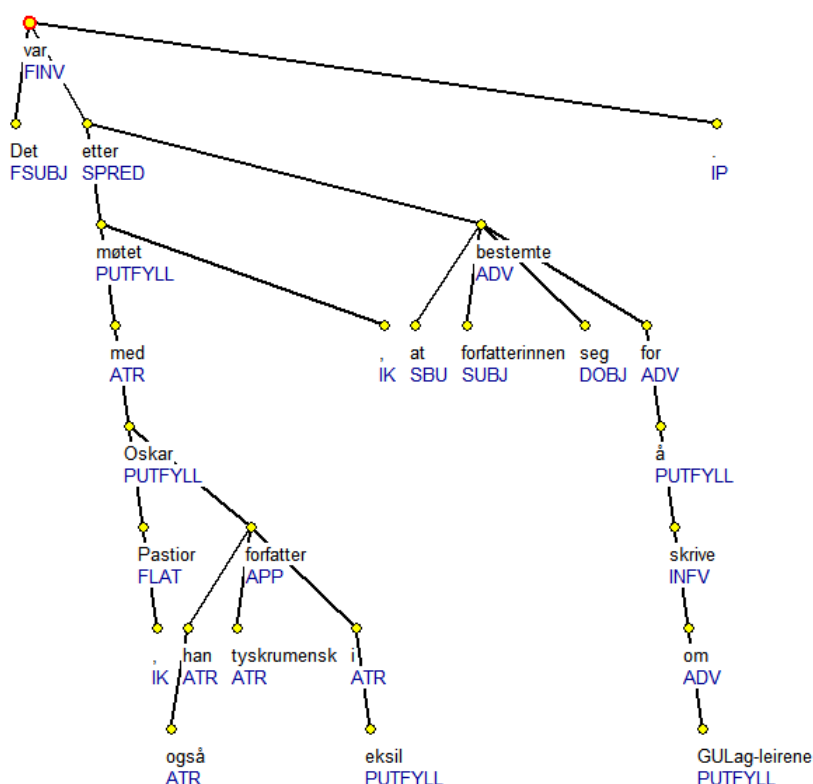
(b)



In another interesting example (4.49 a), we find the delimiting NP *også han*. The context is an article about a female German-Romanian writer in exile, and here, the NP is used to show that *Oskar Pastior* can be added as a subset to the set of people who are German-Romanian writers in exile, to which the female author already belongs.

- (4.49) (a) *Det var etter møtet med Oskar Pastior, også han*
 it was after meeting+the with Oskar Pastior also he
tyskrumensk forfatter i eksil, at forfatterinnen
 German-Romanian writer in exile that female author+the
bestemte seg for å skrive om GULag-leirene.
 decided refl for to write about GULag camps+the
 'It was after the meeting with Oskar Pastior, also a
 German-Romanian writer in exile, that she decided to write about
 the GULag camps.' (Aftenposten 011)

(b)



Finally, it should be mentioned that there is a gradual transition from subset appositions to free predicatives, as the head of a subset apposition does not need to be a noun, but can also be an adjective, a participle, etc.

- (4.50) *Tre båter - den ene hvit, den andre gul og den tredje blå - lå for kai i havna.*
 three boats the one white the second yellow and the third blue lay for quay in harbour+the
 'Three boats, the first white, the second yellow and the third blue, lied alongside the quay in the harbour.'

We will, for the sake of consistency, use a subset apposition analysis also in such cases.

4.3 Dependents on prepositions

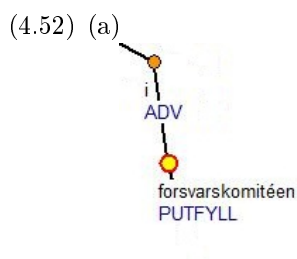
In this section, we explain how we analyze dependents on prepositions. However, we will not explain here what kind of words take prepositional dependents and what functions prepositions should have, as this is described in detail elsewhere (see 3.5, 4.1.2, 4.1.3.3 and 3.8. For the distinction between transitive and intransitive prepositions, see 3.6).

4.3.1 Prepositional complement

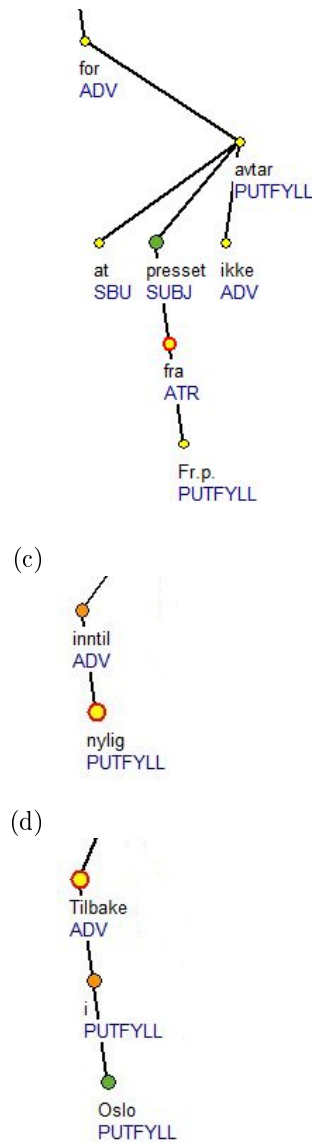
Transitive prepositions can get many types of complements. In (4.51 a) the preposition takes a nominal complement, in (4.51 b) a clause, in (4.51 c) an adverb and in (4.51 d) another preposition. The preposition is made bold in the examples, the complement is emphasized.

- (4.51) (a) *Ap. får også lederverv **i** forsvarskomiteén [...].*
 Ap. gets also leader office in defence committee+the
 'Ap. also gets the leader office in the defence committee'
 (Aftenposten 001)
- (b) *Og Hagen vil sørge **for** at presset fra Fr.p.
 and Hagen will care for that pressure+the from Fr.P.
 ikke avtar.
 not decreases
 'And Hagen will make sure that the pressure from Fr.P. will not
 decrease.' (Aftenposten 001)*
- (c) *Familiebilder som **inntil** nylig har hengt over
 Family pictures which until recently have hung over
 pianoet hjemme i Breivik.
 piano+the home in Breivik
 'Family pictures which, until recently, hung over the piano at home
 in Breivik.' (Aftenposten 003)*
- (d)
***Tilbake** i Oslo er han en sliten mann [...].*
 back in Oslo is he a tired man
 'Back in Oslo, he is a tired man.' (Aftenposten 001)

The complement shall have the function PUTFYLL (*preposisjonsutfylling*, 'prepositional complement').



(b)



Prepositions such as *over*, *rundt* etc. are often used in quantity indications, as in (4.53):

- (4.53) *Det er bra, men ikke bra nok for et land med*
 it is good but not good enough for a country with
over 26 millioner innbyggere.
 over 26 million inhabitants
 ” (Dagbladet 004)

The scope of *over* is not the whole noun phrase: The preposition indicates the amount, that the number of inhabitants exceeds 26 million. Therefore, only the quantity indication should be PUTFYLL. In the case of complex number expressions with a determiner and a number noun, as in (4.53), we have chosen to let only the determiner be PUTFYLL. The analysis will thus look like this:

(4.54)



4.3.2 Postponed prepositions

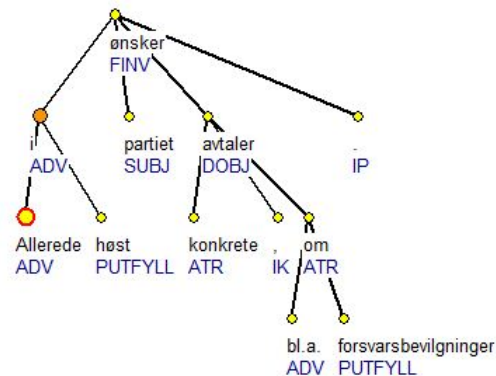
Some prepositions can follow their complement (see Faarlund et al. 1997, 463-464). This occurs particularly with when *over* and *rundt*, as in (4.55 a) and (4.55 b):

- (4.55) (a) [...] *han er fotografen som flyr verden over* [...].
 he is photographer+the who flies world over
 'He is the photographer who flies all over the world.' (Dagbladet 001)
- (b) *Forrige helg var det premiere Norge rundt på*
 last weekend was it premiere Norway around on
 filmen deUSYNLIGE.
 film+the deUSYNLIGE
 'Last weekend, the film deUSYNLIGE had its premiere all around
 Norway.' (Dagbladet 001)

4.3.3 ADV on prepositions

Prepositions can be modified by adverbs. In such cases, the adverb is ADV-dependent on the preposition:

- (4.56) (a) *Allerede i høst ønsker partiet konkrete avtaler, bl.a.*
 already in fall wished party+the concrete deals e.g.
 om forsvarsbevilgninger.
 about defence fundings
 'Already this fall, the party wished concrete deals, e.g. concerning
 defence fundings.' (Aftenposten 001)
- (b)



See also 3.5.2.

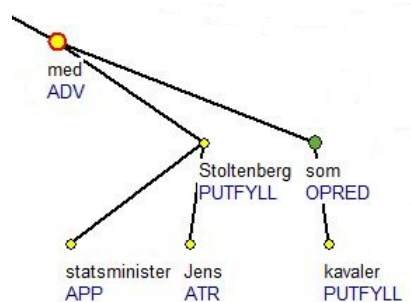
4.3.4 Predicative with *med*

With the preposition *med*, 'with', and probably also *uten*, 'without', there is often a kind of predicative to the prepositional complement. In (4.57), there are good reasons to consider 4.57 as a predicative to *Jens Stoltenberg*:

- (4.57) *Prinsesse Märtha Louise savnet neppe sin Ari med statsminister Jens Stoltenberg som kavalier [...].*
 Princess Märtha Louise missed hardly her Ari with Prime Minister Jens Stoltenberg as partner
 'Princess Märtha Louise hardly missed her Ari with Prime Minister Jens Stoltenberg as partner.' (Aftenposten 001)

We let the predicative phrase be OPRED on the preposition, as the analysis (4.58) shows. See also 3.8.3 for more on this.

(4.58)

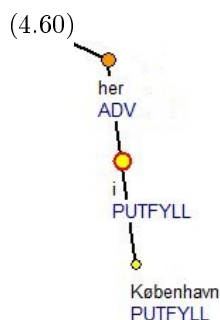


4.3.5 New prepositions

In Faarlund et al. 1997, a number of words which previously were classified as adverbs, are listed as prepositions. These include *her*, 'here', *der*, 'there', *hit*,

'over here' and *dit*, 'over there'. These are mostly intransitive. In a sentence like (4.59), however, we will count the preposition as transitive, as the analysis in (4.60) shows:

- (4.59) *Tekster og musikken spiller en større rolle her i*
 texts and music+the plays a larger role here in
 København.
 Copenhagen
 'The texts and the music plays a larger role here in Copenhagen.'
 (Dagbladet 003)



Nord, 'north', *sør* 'south' etc. are also counted as prepositions, cf. 1.1.5.

4.4 Coordination

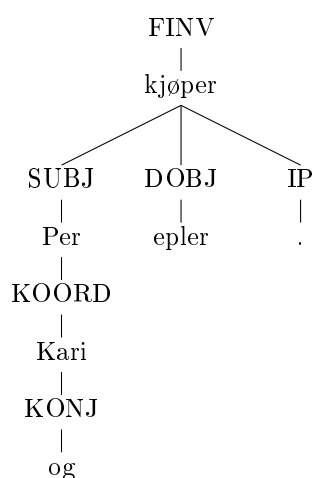
As we mentioned in section 2.1, coordination is a problem in Dependency Grammar: Dependencies are by definition asymmetrical, they show the relationship between a superordinate and a subordinate element; coordination, however, seems to imply a symmetrical relationship between two or more conjuncts. In (4.61), there is no obvious asymmetric relationship between *Per* and *Kari*:

- (4.61) *Per og Kari kjøper epler.*
 Per and Kari buy apples
 'Per and Kary buy apples'

Coordination could have been represented by allowing the conjunction to be head and let the conjuncts be dependent on the conjunction. There are several problems with this solution: firstly, it will not work if the conjunction is missing, secondly, the conjunction could then be associated with virtually all syntactic functions, and it would therefore be more difficult to search in the corpus.

Although coordination appears to be symmetric, we have found it more appropriate to let the first conjunct always be the head, and let other conjuncts be dependent on the first conjunct with function KOORD. The conjunction is dependent on the head of the nearest conjunct to the right, which in most cases will be the last conjunct, with the function KONJ (this analysis is partly based on Nivre et al. 2006). The analysis of (4.61) will in other words look like this:

(4.62)



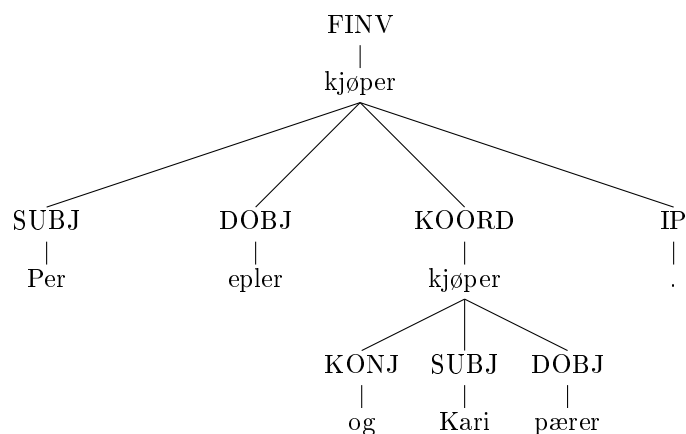
For the placement of punctuation in coordination, see sub-chapter 5.2.

Sentence coordination

Coordination of sentences, as in (4.63), is treated in exactly the same way, as the tree in (4.64) shows. The verb in the second conjunct is dependent on the verb of the matrix clause, the conjunction is on the verb of the second conjunct:

- (4.63) *Per kjøper epler og Kari kjøper pærer.*
 Per buys apples and Kari buys pears
 'Per buys apples and Kari buys pears.'

(4.64)



Occasionally a sentence consists of two relatively independent sentences which are not separated by a conjunction, such as (4.65). We analyze such cases as sentence coordination.

- (4.65) *Fremgangsmåten er enkel, det er bare å legge plata i*
 method+the is simple it is just to lay plate+the in
sekken før du sender den gjennom røntgenmaskinen.
 sack+the before you send it through x-ray machine+the

'The method is simple, you just put the plate in the sack before you send it through the x-ray machine.' (Dagbladet 001)

Shared constituents

Often, one or more constituents are shared by all conjuncts, as *i butikken* in (4.66). We have chosen not to represent that an element is shared, but let such constituents be dependent on the closest conjunct. *I butikken* in (4.66) will therefore be analyzed as part of the second conjunct.

- (4.66) *Per kjøper epler og Kari kjøper pærer i butikken.*
 Per buys and Kari buys pears in the shop
 'Per buys apples and Kari buys pears in the shop.'

The conjuncts should fill the same function in the sentences

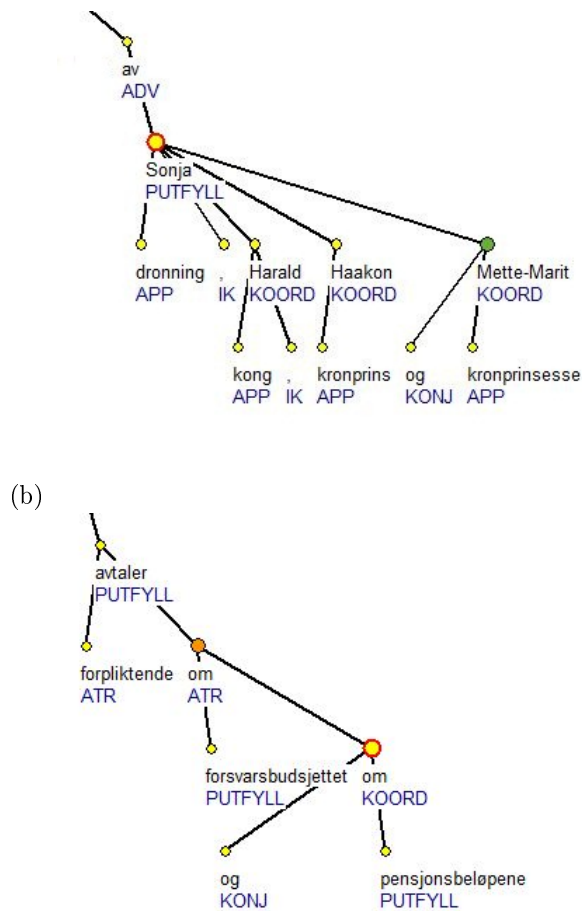
An important principle in our coordination analysis is that the two conjunct should be able to fill the same function in the sentence³. Stated another way, each conjunct should be able to have the same syntactic function as first conjunct. The following examples illustrate this:

- (4.67) (a) *De ble ønsket velkommen av dronning Sonja,*
 they became wished welcome of Queen Sonja
kong Harald, kronprins Haakon og kronprinsesse
 King Harald Crown Prince Haakon and Crown Princess
Mette-Marit.
 Mette-Marit
 'They were welcomed by Queen Sonja, King Harald, Crown Prince Haakon and Crown Princess Mette-Marit.' (Aftenposten 001)
- (b) *Hagen er særlig interessert i forpliktende avtaler*
 Hagen is particularly interested in binding deals
om forsvarsbudsjettet og om pensjonsbeløpet.
 about defence budget+the and about pension amount+the
 'Hagen is particularly interested in binding deals on the defence budget and on the size of the pension pot.' (Aftenposten 001)

In (4.67a), all the conjuncts can individually function as PUTFYLL to *av*, but they cannot take the function of *av*. Therefore, we have coordination of PUTFYLL in (4.67a), as (4.68a) shows. In (4.67b), however, the attributive preposition *om* repeated, and the second conjunct alone cannot be PUTFYLL. Therefore, the second conjunct must be coordinated with the preposition, cf. (4.68b).

- (4.68) (a)

³The requirement that the conjunct should be able to have the same function is not absolutely mandatory. If there are good reasons to believe that two phrases with different functions are coordinated with each other, it should of course be reflected in the analysis.



Quite often, one or more words are implied or elided in one of the conjunct, as in (4.69 a) and (4.69 b).

- (4.69) (a) *Selv om Bondevik [...] gir en del vage svar, og*
 even if Bondevik gives a part vague answers and
også noen negative, er tonen i svarbrevet
 also some negative is tone+the in answer letter+the
gjennomgående imøtekommende.
 generally obliging
 'Even though Bondevik gives some vague answers, and also some
 negative ones' (Aftenposten 001)

- (b) *Han ble boende hos faren, storesøstera hos*
 he became living with father+the big sister+the with
mora.
 mother+the
 'He continued living with his father, the big sister with her mother.'
 (Dagbladet 001)

In such cases it is, at least apparently, difficult to follow the principle that the two conjuncts should be able to have the same function. In 5.1.3, we show how

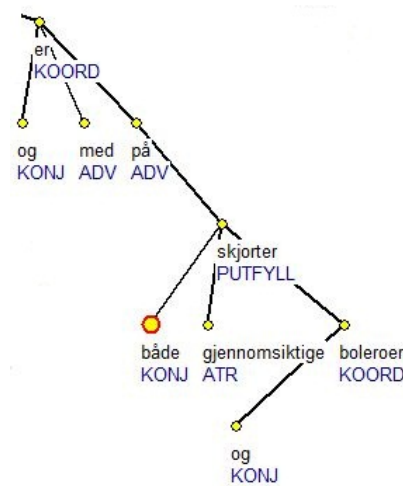
such structures should be analyzed.

4.4.1 Coordination with two conjunctions

In the case of coordination with two conjunctions, such as *både...og*, 'both...and', *enten...eller*, 'either...or', and the like, the first conjunction shall be dependent on the first conjunct. The function is KONJ:

- (4.70) (a) *Nå skjønner han at det er sport og en rolle,*
 now understands he that it is sport and a role
og er med på både gjennomsluktige skjorter og boleroer.
 and is with on both transparent shirts and boleros
 'Now, he understands that it is sport and a role, and he accepts both
 transparent shirts and boleros.' (Dagbladet 002)

(b)

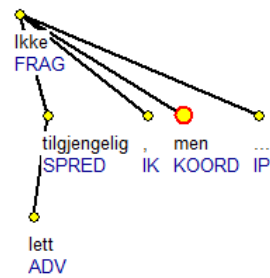


4.4.2 Coordination without conjuncts

Sometimes one can find instances of conjunctions used in the end of a sentence, without a conjunct following it. In such cases, we analyze the conjunction himself as head of the "conjunct" - the conjunction takes the place an ordinary conjunct would have had, and gets the function KOORD. The analysis is thus parallel to the analysis of ellipses (5.1). In (4.71 a), we see how this turns out in practice:

- (4.71) (a) *Ikke lett tilgjengelig, men ...*
 not easily accessible, but
 'Not, easily accessible, but ...' (Aftenposten 003)

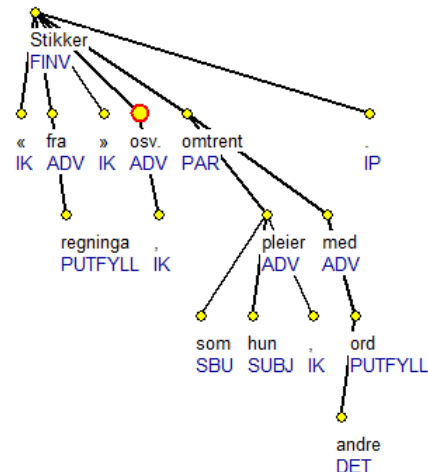
(b)



Abbreviations of type *osv.* and *etc.* can semantically be viewed as empty conjuncts, but as *osv.* and *etc.* are not conjunctions in general sense, we decided to give them another analysis. In the morphology, we tag them as adverbs, and in the syntax, we make them dependents on the head of the last preceding conjunct (i.e. if *osv.* or *etc.* is preceded by a series of conjuncts, as in *Her på bakeriet selger vi boller, osv.*, 'Here at the bakery, we sell buns, donuts, pretzels, etc.'). If they are not preceded by a series of conjuncts (as in *Her på bakeriet selger vi boller, osv.*, 'Here at the bakery, we sell buns etc.'), *osv.* and *etc.* get the function ADV based on their morphological *adv*-tag, but this is replaced by ATR if *osv.* or *etc.* is dependent on a noun (cf. 4.1.2).

- (4.72) (a) “Stikker fra regninga osv.”, omtrent som hun
 run from bill+the etc. approximately as she
pleier, med andre ord.
 usually-does with other words
 “Runs from the vill, etc”, more or less as she usually does, in other
 words.’ (Aftenposten 009)

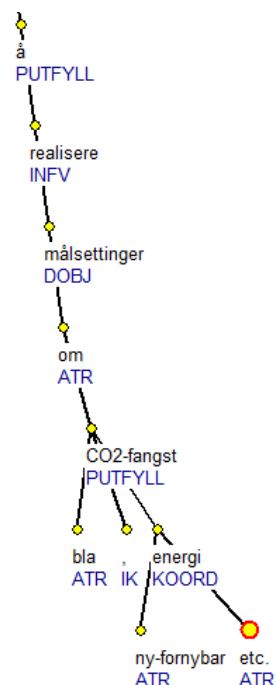
(b)



- (c) [...] å realisere målsettinger om bla CO2-fangst,
 to realize goals about e.g. CO2 capture
 ny-fornybar energi etc.
 new-renewable energy etc.

'[...] to realize goals for CO2 capture, new-renewable energy etc.'
(Aftenposten 004)

(d)



4.4.3 *Enten-eller* coordination as adverbial clauses

A sentence containing coordination with *enten* and *eller* may function as an adverbial clause in the superordinate sentence:

- (4.73) *Du spiser maten din enten du vil eller ikke!*
you eat food+the your either you want or not
'You will eat your food, whether you want to or not.'

- (4.74) *Vi er alle like enten vi har svart eller hvit hud.*
we are all alike either we have black or white skin
'We are all alike, whether we have black or white skin.'

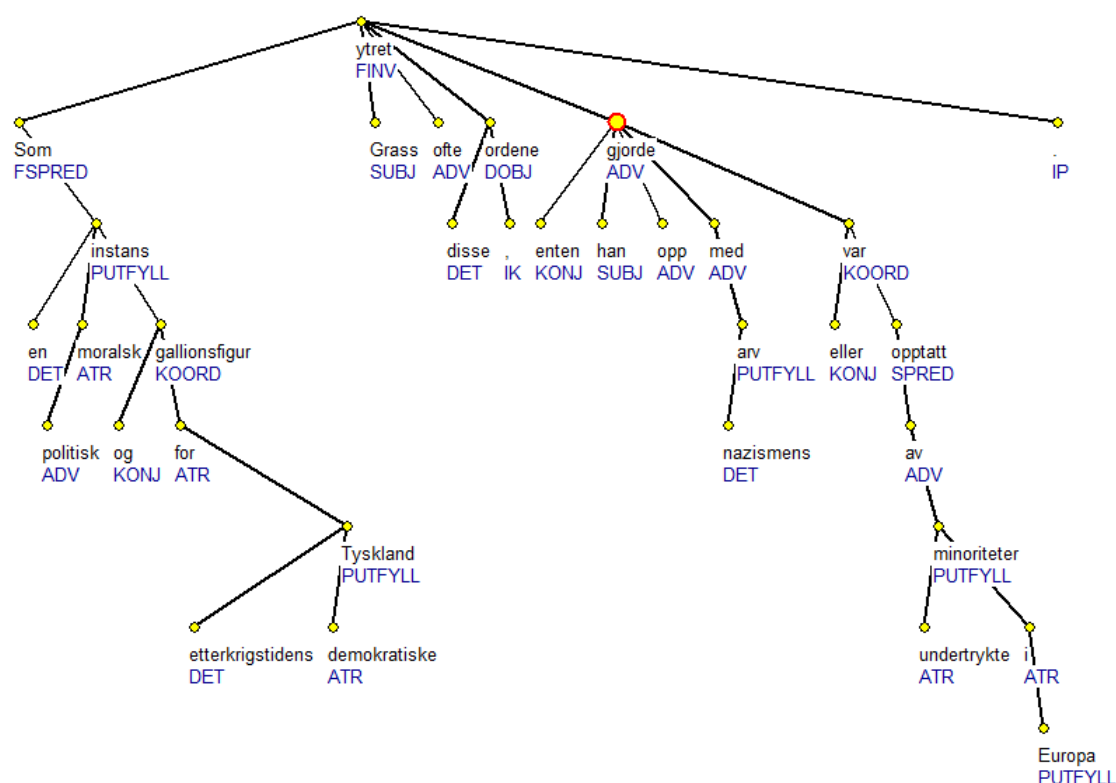
Note that this clearly is a case of subordination, as independent *enten-eller*-coordinated sentences always have subject-verb inversion (cf. 4.75), which these have not. Fronting of an *enten-eller* clause also leads to subject-verb inversion in the matrix clause (cf. 4.76).

- (4.75) *Enten har de svart, eller så har de hvit hud.*
either have they black, or so have they white skin
'Either they have black or they have white skin.'

- (4.76) *Enten vi har svart eller hvit hud, (så) er vi alle like.*
either we have black or white skin so are we all alike
'Whether we have black or white skin, we are all alike.'

Still, such subordinate clauses do not contain a complementizer. It is therefore tempting to analyze *enten* as the complementizer in these cases, and there may be good linguistic reasons for believing that a grammaticalization of *enten* of this kind has taken place. However, to ensure a common morphological and syntactic analysis of all cases of *enten*, we have instead chosen to analyze these subordinate clauses as clauses lacking a complementizer, which is not an uncommon phenomenon either. *Enten* and *eller* get a regular conjunction analysis inside the clause, while the whole clause - specifically the finite verb that functions as its head - will get the function ADV, and becomes an adverbial dependent in the superordinate clause:

- (4.77) (a) *Som en politisk moralsk instans og gallionsfigur for*
 as a political moral authority and figurehead for
etterkrigstidens demokratiske Tyskland ytret Grass
 post-war period+the's democratic Germany uttered Grass
ofte disse ordene, enten han gjorde opp med
 often these words+the either he did up with
nazismens arv eller var opptatt av undertrykte
 Nazism+the's legacy or was preoccupied of oppressed
minoriteter i Europa.
 minorities in Europe
 'As a political moral authority and a figurehead for the democratic
 Germany of the post-war period, Grass often uttered these words,
 whether he confronted the legacy of Nazism or was preoccupied with
 oppressed minorities in Europe.' (Aftenposten 011)
- (b)



4.4.4 Adverbs and adverbial phrases as quasi-conjunctions

Certain adverbs and fixed adverbial expressions often seem to behave like conjunctions. This applies to *eventuelt*, 'possibly', *langt mindre*, 'much less', *dels*, 'partly' etc.

- (4.78) *Vi reiser med fly, eventuelt med tog.*
 we travel with plane possibly with train
 'We travel by plane, possibly by train.'

- (4.79) *Dei materielle føresegnene i lova regulerer dels i*
 the material regulations in law+the regulate partly in
kva situasjonar eit statsorgan skal bruka anten nynorsk
 what situations a state agency shall use either Nynorsk
eller bokmål, dels korleis bruksfordelinga mellom bokmål
 or Bokmål partly how use distribution+the between Bokmål
og nynorsk skal vera for allment dokumenttilfang som
 and Nynorsk shall be for common document material which
ikkje kjem inn under dei spesifikke føresegnene.
 not comes in under the specific regulations

'The material regulations in the law regulate partly in what situations a state agency should use either Nynorsk or Bokmål, partly how the distribution of use should be between Bokmål and Nynorsk for common

document material which is not covered by the specific regulations.’
(Mål og meining 003)

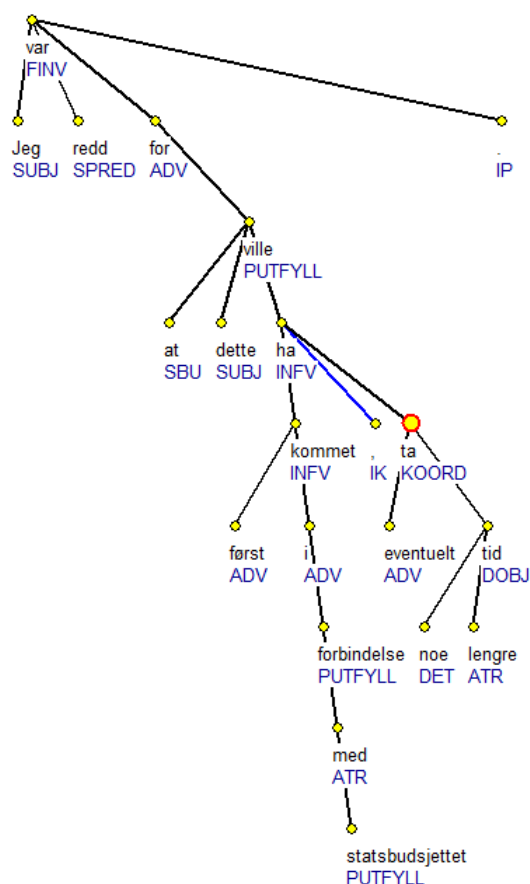
Note, however, that in all these cases, you can supplement the example with a real conjunction:

- (4.80) *Vi reiser med fly, eller eventuelt med tog.*
we travel with plane or possibly with train
‘We travel by plane, or possibly by train.’

- (4.81) *Dei materielle føresegnene i lova regulerer dels i*
the material regulations in law+the regulate partly in
kva situasjonar eit statsorgan skal bruka anten nynorsk
what situations a state agency shall use either Nynorsk
eller bokmål, og dels korleis bruksfordelinga mellom
or Bokmål and partly how use distribution+the between
bokmål og nynorsk skal vera for allment dokumenttilfang
Bokmål and Nynorsk shall be for common document material
som ikkje kjem inn under dei spesifikke føresegnene.
which not comes in under the specific regulations
‘The material regulations in the law regulate partly in what situations a
state agency should use either Nynorsk or Bokmål, and partly how the
distribution of use should be between Bokmål and Nynorsk for common
document material which is not covered by the specific regulations.’
(Mål og meining 003)

We therefore choose always to analyze these elements as regular adverbials in both the morphology and the syntax. If no real conjunction occurs in the sentence, the last conjunct analyzed is analyzed as lacking a CONJ-element.

- (4.82) (a) *Jeg var redd for at dette først ville ha kommet*
I was afraid for that this first would have come
i forbindelse med statsbudsjettet, eventuelt ta noe
in connection with state budget+the possibly take some
lengre tid.
longer time
‘I was afraid that this first would have come in connection with the
state budget, possibly take somewhat more time.’ (Stortinget 001)
- (b)



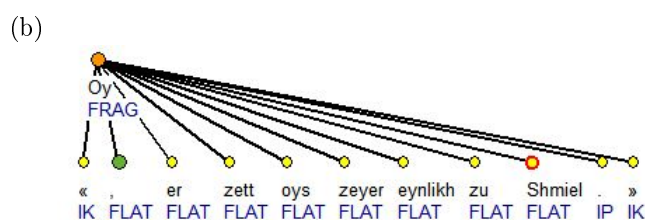
4.5 Flat structure

In some cases, it is not appropriate or possible to give a hierarchical syntactic structure to a construction, such as for quotes in foreign languages and names. Such constructions could have been merged into a single token, but since it is important for some users of the corpus not to have tokens containing whitespaces, we have chosen to show this in the syntactic annotation instead. Such constructions will have what we call a flat structure: We set the first or last word in the construction as the head, and let all the other tokens in the construction be dependent on the head with the function FLAT. Tokens with the function FLAT do not have dependents themselves, all dependents must be to the head of the flat structure.

4.5.1 Quotes in foreign languages

Quotes in foreign languages shall have a flat structure, with the first word as head:

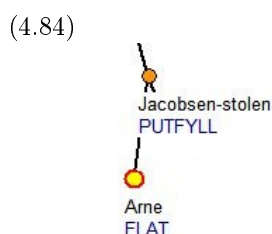
- (4.83) (a) «*Oy, er zett oys zeyer eynlikh zu Shmiel.*» [Quote in Yiddish]
(Aftenposten 012)



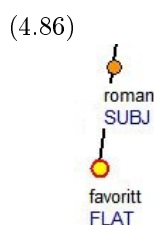
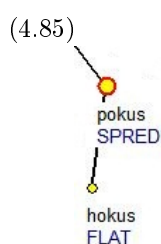
See also 1.2.3.1.

4.5.2 Complex expressions with whitespace, *au pair* og non-standard word division

Complex expressions such as *Arne Jacobsen-stolen*, 'Arne Jacobsen chair+the', function, syntactically speaking, as a single word, and they therefore get a flat structure in NDT. In such complex expressions, it is the last word that carries the relevant part of speech, inflection etc. We therefore let the last word be the head. The analysis of the *Arne Jacobsen-stolen* is therefore as follows (from Aftenposten 009):



Words such as *au pair* and *hokus pokus*, 'hocus pocus', are analyzed the same way, see (4.85) (Dagbladet 004), and likewise words with non-standard word division, see the analysis of the *favoritt roman*, 'favorite novel' in (4.86) (from Dagbladet 003):

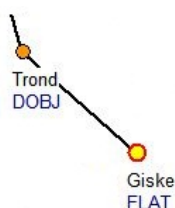


See also 1.2.3.2 and 1.2.6.1.

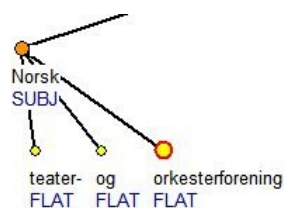
4.5.3 Proper nouns

Proper nouns consisting of several words should also have a flat structure in NDT. In many cases, all words have capitalized initials, as in *Trond Giske*. In other cases, the first word has a capitalized initial, while some or all of the other words have a non-capitalized initials, as in *Norsk teater- og orkesterforening* and *Forum for Kultur og Næringsliv*. In the morphology, capitalized words are tagged as proper nouns, whereas non-capitalized words get their common morphological tags (cf. 1.1.6). Since the first word in such proper names is always tagged as a noun in the morphology, we have chosen to let it be the head in the syntax, see (4.87 a), (4.87 b) and (4.87 c) (all from Aftenposten 012):

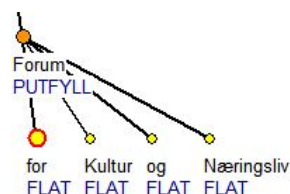
(4.87) (a)



(b)



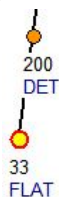
(c)



4.5.4 Numbers with whitespaces

Numbers written with whitespaces, like 33 200, shall have a flat structure. The last part of the number will be the head, see (4.88) (from Dagbladet 014):

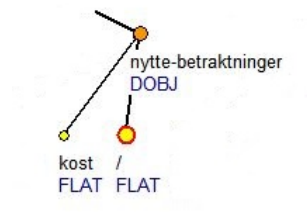
(4.88)



4.5.5 Complex words with a slash

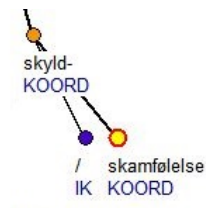
The first element, the last element and the slash in compounds such as *kost/nytte-betraktning*, 'cost/benefit considerations' (Aftenposten 011) are tokenized separately, cf. 1.2.5.3. In such cases, the last element is the head, while the first element and the slash are FLAT-dependents:

(4.89)



If the first element contains a hyphen indicating that parts of the word is omitted, as *skyld-/skamfølelse*, 'guilt/shame feeling' (Aftenposten 007), we consider it as parallel to *skyld- og skamfølelse*. Therefore, we choose a regular coordination analysis, cf. 1.2.4:

(4.90)



Kapittel 5

Other questions

5.1 Ellipsis

It often happens that a word or paragraph is implied in a sentence. This applies especially to verbs and nouns, as in the following examples:

- (5.1) (a) *Per har et rødt og et grønt eple. Han gir bort det røde.*
Per has a red and a green apple he gives away the red
'Per has a red and a green apple. He gives away the green one.'
- (b) *Aldri en dag uten at passet er med.*
never a day without that passport+the is with
'Never a day without bringing the passport'(Dagbladet 001)
- (c) *Kari kjøper epler og Per plommer.*
Kari buys apples and Per plums
'Kari buys apples and Per plums'

In the second sentence of (5.1 a) the noun *eple* is implied from the preceding sentence. (5.1 b) lacks a finite verb in the superordinate clause, but the adverbial phrases *aldri* and *uten at passet er med* shows that we must imply a verb of some kind, e.g. *går*, 'goes'. In (5.1 c), we have two coordinated constituents. Both have both subject and object, and must therefore be regarded as finite sentences, but in the second conjunct, the verb is implied. We will hereafter refer to such omitted words and phrases as *ellipses*. The omitted word/phrase is said to be *elided*.

We cannot use empty nodes in this corpus. Therefore, we represent ellipses in a different way. In the following subsections we will cover how we treat noun ellipses (5.1.1), elliptical matrix clauses in (5.1.2) and ellipses and coordination (5.1.3).

5.1.1 Noun ellipsis

Adjectives to nouns

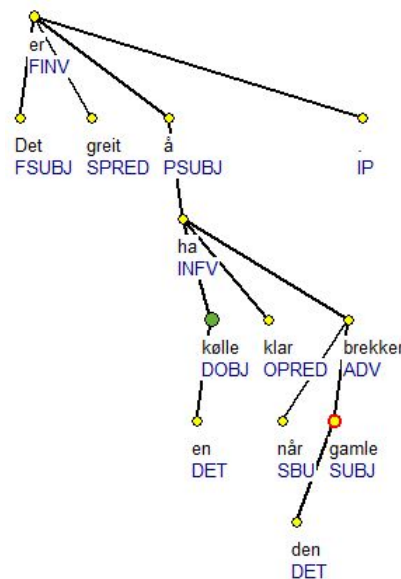
It occurs quite often that only an adjective, often in combination with a determiner, occurs where a noun is expected. Sometimes the noun implied, for

example because it occurs earlier in the sentence, as in (5.2 a), or the noun is nominalized, as in (5.2 b).

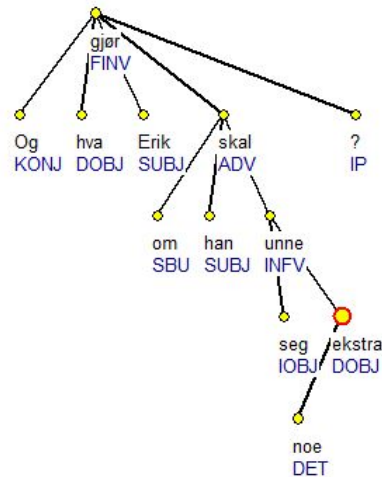
- (5.2) (a) *Det er greit å ha en kølle klar når **den gamle** brekker.*
 it is ok to have a baton ready when the old one breaks.
 breaks
 'It is ok to have a baton ready when the old one breaks.' (Dagbladet 001)
- (b) *Og hva gjør Erik når han skal unne seg **noe ekstra?***
 and what does Erik when he shall allow refl something extra
 extra
 'And what does Erik do when he allows himself something extra' (Dagbladet 001)

In cases where you have both the determiner and an adjective, as in (5.2 a) and (5.2 b), we let be the adjective head. While the adjective normally only take ADV-dependent (cf. 3.5.2), the adjective in such elliptical constructions will in addition be able to take typical nominal dependent, as DET and ATR. The determiner and under (5.2 a) and (5.2 b) is therefore DET on the adjective:

- (5.3) (a)



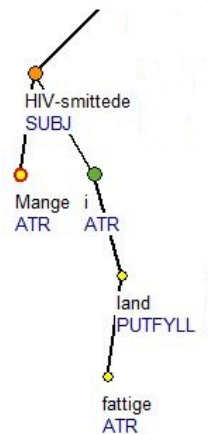
- (b)



The subject of (5.4 a) is a nominalized adjective, *HIV-smittede*, taking ATR-dependent, see the analysis in (5.4 b):

- (5.4) (a) *Mange HIV-smittede i fattige land får ikke medisiner som kan gi dem et bedre og lengre liv [...].*
 many HIV infected in poor countries get nor medicines
som kan gi dem et bedre og lengre liv [...].
 which can give them a better and longer life
 'Many HIV positive in poor countries do not get medicines which can give them a better and longer life.' (Aftenposten 001)

(b)



Other attributes than adjectives, e.g. prepositions, can sometimes take the place of an elided noun. We follow the same procedure also in such cases.

Determiner for noun

Sometimes, a determiner occurs without a noun or adjective. This is true for numerals, which are determiners in the morphological annotation, but which often function as nouns:

- (5.5) (a) *I 1998 debuterte han som spillefilmregissør [...].*
 in 1998 made-debut he as feature film director
 'In 1998, he made his debut as feature film director.' (Dagbladet 001)
- (b) *Rundt 1200 stilte seg bak.*
 around 1200 place refl behind
 'Around 1200 stood behind it.' (Aftenposten 001)

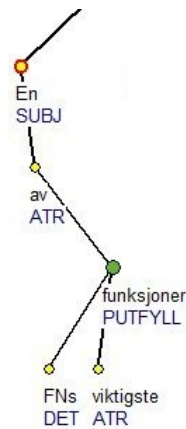
In such cases, the determiner becomes the head.

A determiner is the head also when we have partitive constructions with the preposition *av*, cf. 4.1.2.2:

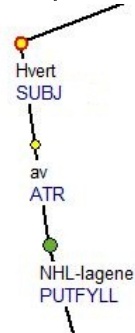
- (5.6) (a) *En av FNs viktigste funksjoner er imidlertid som*
 one of UN's important.sup functions is however as
forum for diskusjon og møtested for mennesker fra
 forum for discussion and meeting place for people from
hele verden [...].
 whole world
 'One of the UN's most important functions, however, is being a forum for discussion and a meeting place for people from all over the world.' (Aftenposten 001)
- (b) *Hvert av NHL-lagene som spilte i Europa for to*
 each of NHL-teams+the which played in Europe for two
uker siden fløy over cirka 65 personer og store
 weeks ago flew over circa 65 people and large
mengder utstyr.
 quantities equipment
 'Each of the NHL teams which played in Europe two weeks ago, flew over ca. 65 people and large quantities of equipment.' (Dagbladet 001)

Words such as *en* in (5.6 a) and *hver* in (5.6 b) get *determiner* as part of speech in the morphological annotation. The preposition *av* will be ATR on the determiner:

- (5.7) (a) Analysis of 5.6 a:



(b) Analysis of 5.6 b:



5.1.2 Elliptical matrix clauses

In the case of matrix clauses without a finite verb, as (5.8 a) and (5.8 b), we use the function FRAG (see sub-chapter 3.1.3):

(5.8) (a) *Klar beskjed:*

clear message

'Clear message:' (Dagbladet 001)

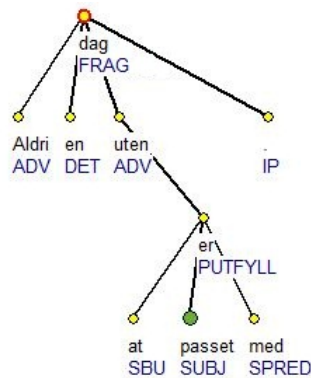
(b) *Aldri en dag uten at passet er med.*

never a day without that passport+the is with

'Never a day without bringing the passport' (Dagbladet 001)

We do not assume that sentences without a finite verb are necessarily elliptical. Sometimes, however, the head of FRAG-constructions takes dependents typically associated with verbs. For example, it is natural to imply a verb in (5.8 b), such as *går*, 'goes', taking *uten* as an adverbial. In such cases, we let one of the elided verb's arguments to be the head with the function FRAG. In (5.8 b), *dag* will be the head. *Aldri* and *uten*, which probably are adverbials to the elided verb, become ADV-dependents to *dag*:

(5.9) Analysis of 5.8 b:



Verbs can take many types the dependents, and it is therefore difficult to decide which word should replace the verb as head. We've created a prominence hierarchy for arguments to verbs, shown in (5.10)¹. The further to the left in the hierarchy, the more prominent the argument is:

- (5.10) non-finite verb > subject > indirect object, subject predicative > potential subject, direct object > object predicative > adverbial

In the case of verb ellipsis, the most prominent argument in this hierarchy should be the head. In the case of noun ellipsis, we consider adjectives as more prominent than determiners, cf. 5.1.1.

5.1.3 Ellipsis and coordination

Our coordination analysis requires that all conjuncts fill the same grammatical function in the sentence (cf. subsection 4.4). Very frequently, however, the head of one or more of the conjuncts is implied, cf (5.11 a) and (5.11 b).

- (5.11) (a) *Selv om Bondevik [...] gir en del vage svar, og*
 even if Bondevik gives a part vague answers and
også noen negative, er tonen i svarbrevet
 also some negative is tone+the in answer letter+the
gjennomgående imøtekommende.
 generally obliging
 'Even though Bondevik gives some vague answers, and also some negative ones' (Aftenposten 001)
- (b) *Han ble boende hos faren, storesøstera hos*
 he became living with father+the big sister+the with
mora.
 mother+the
 'He continued living with his father, the big sister with her mother.'
 (Dagbladet 001)

In (5.11 a), there are good reasons to consider *negative* as coordinated with *svar*, e.g. because we have a determiner in the second conjunct. In the second

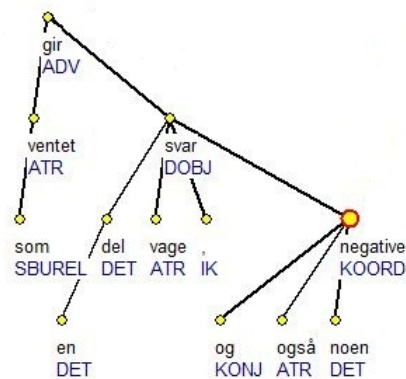
¹This hierarchy is created on the basis of binding and control relationships which exist between different types of arguments: The most prominent arguments can be binders and controllers of less prominent arguments, but not vice versa.

conjunct in (5.11 b), we must assume that *ble boende* is elided, because we have the adverbial prepositional phrase *hos mora*. In this corpus, the analysis of coordination with noun ellipsis and coordination with verb ellipsis is treated slightly differently.

5.1.3.1 Coordination with noun ellipsis

As we showed in section 5.1.1, ellipsis of nouns is frequent independently of coordination. In other words, *noen negative* in (5.11 a) could have occurred as object also in a sentence with no coordination. Therefore, the requirement that the two conjunct can fill the same function, is fulfilled. Therefore, we can let the adjective in the second conjunct be coordinated with the noun in the first, without creating any problems:

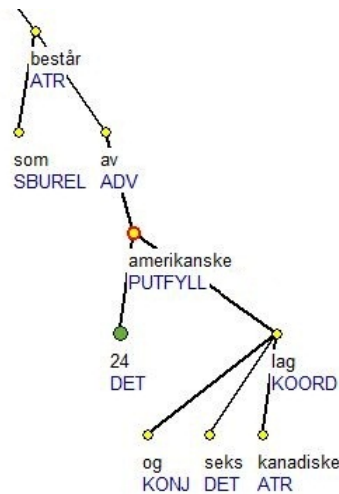
(5.12)



Note that the ellipsis occasionally occurs in the first conjunct, as in the following example:

- (5.13) (a) *Det er andre gang National Hockey League, verdens beste liga, som består av 24 amerikanske og seks kanadiske lag, serieåpner i Europa.*
 it is second time National Hockey League worlds best league which consists of 24 American and six Canadian teams league-opens in Europe
 'It is the second time National Hockey League, the world's best league, which consists of 24 American and six Canadian teams, opens in Europe.' (Dagbladet 001)

(b) Analysis of (5.13 a):



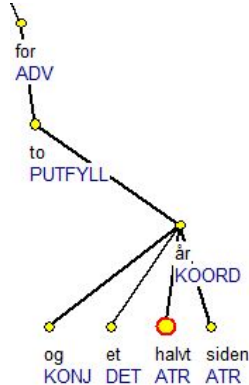
In quantity indications of the type *to og et halvt år*, 'two and a half years', the first conjunct will also be elliptical in our analyses:

- (5.14) (a) *Faren hennes [...] døde for to og et halvt år siden.*
 father+the her died for two and a half year ago

ago

'Her father died two and a half years ago.' (Dagbladet 003)

(b)



5.1.3.2 Coordination with verb ellipsis

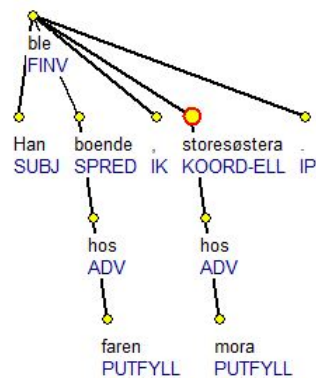
Verb ellipsis, as in (5.11 b), repeated in (5.15), represents a more real threat to the rule that the two conjuncts should be able to have the same function. With the exception of elliptical matrix clauses, it is generally not possible to drop a verb in Norwegian.

- (5.15) *Han ble boende hos faren, storesøstera hos*
 he became living with father+the big sister+the with
mora.
 mother+the

'He continued living with his father, the big sister with her mother.'
(Dagbladet 001)

If we coordinate *storesøstera* with *ble* in (5.15) with the function KOORD, we risk that parsers trained on the corpus will be unable to figure out what the proper use of the KOORD function is. With coordination with verb ellipsis, we therefore use a special function: KOORD-ELL. The analysis of (5.15) is therefore as follows:

(5.16)



The function KOORD-ELL will be used both with ellipsis of finite and non-finite verbs.

We use the prominence hierarchy in 5.10 to determine what should be the head in the conjunct with verb ellipsis. Sometimes, however, more than one word is elided, as in (5.17 a). In (5.17 b), the omitted words are written in brackets:

- (5.17) (a) *FN har også krigsdomstolen for det tidligere Jugoslavia i samme by og for Rwanda i Arusha i Tanzania.*
UN has also war tribunal+the for the former Yugoslavia in same city and for Rwanda in Arusha in Tanzania.

Tanzania

'The UN also has the war tribunal for former Yugoslavia in the same city and for Rwanda in Arusha in Tanzania.' (Aftenposten 001)

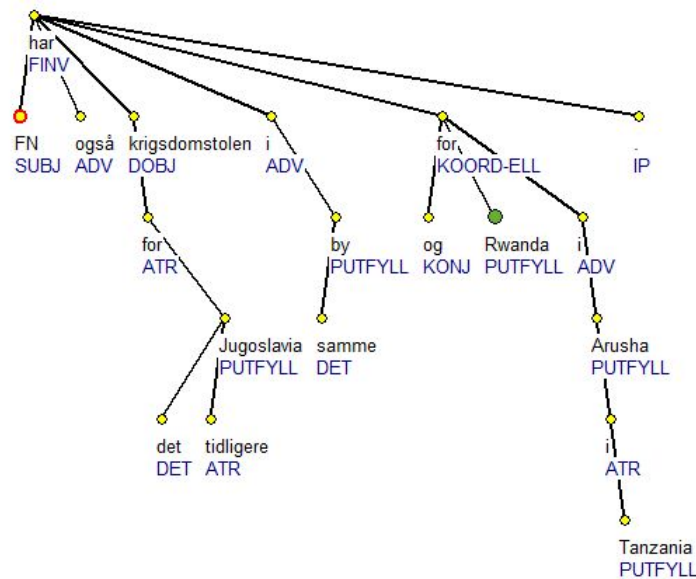
- (b) *FN har også krigsdomstolen for det tidligere Jugoslavia i samme by og [har] [krigsdomstolen] for Rwanda i Arusha i Tanzania.*
UN has also war tribunals for the former Yugoslavia in same city and has war tribunal+the for Rwanda in Arusha in Tanzania.

'The UN also has the war tribunal for former Yugoslavia in the same city and has the war tribunal for Rwanda in Arusha in Tanzania.'

(Aftenposten 001)

As a consequence of our analysis of noun ellipsis, the attributive preposition, *for*, takes the object's position in the second conjunct. The object, in turn, moves up to take the verb's position. *For* is thus head in the second conjunct:

(5.18)

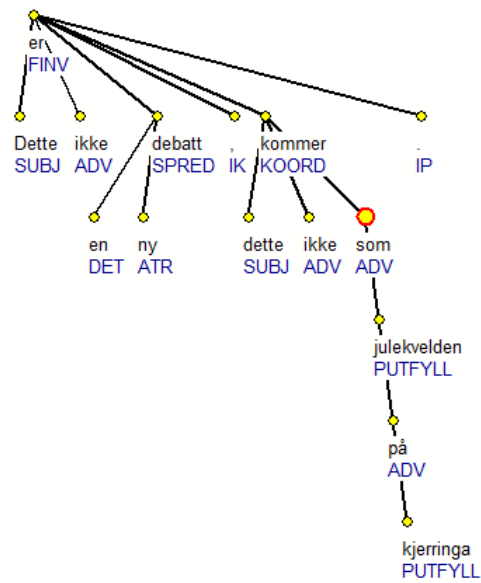


5.1.4 Ellipsis and the prepositions *som* and *enn*

In the case of *som*, 'as', and *enn*, 'than', (as prepositions) there is from time to time an elided verb in the prepositional complement. We will analyze them as we analyze the coordination ellipses: The most prominent constituent in the prominence hierarchy (5.10), is the head of the elliptical clause, and is therefore directly dependent on *som/enn* with the function PUTFYLL. The remaining parts of the clause are dependents on the PUTFYLL-node and get the functions they would have had if the sentence weren't elliptical. Below we provide examples with both *som* and *enn*:

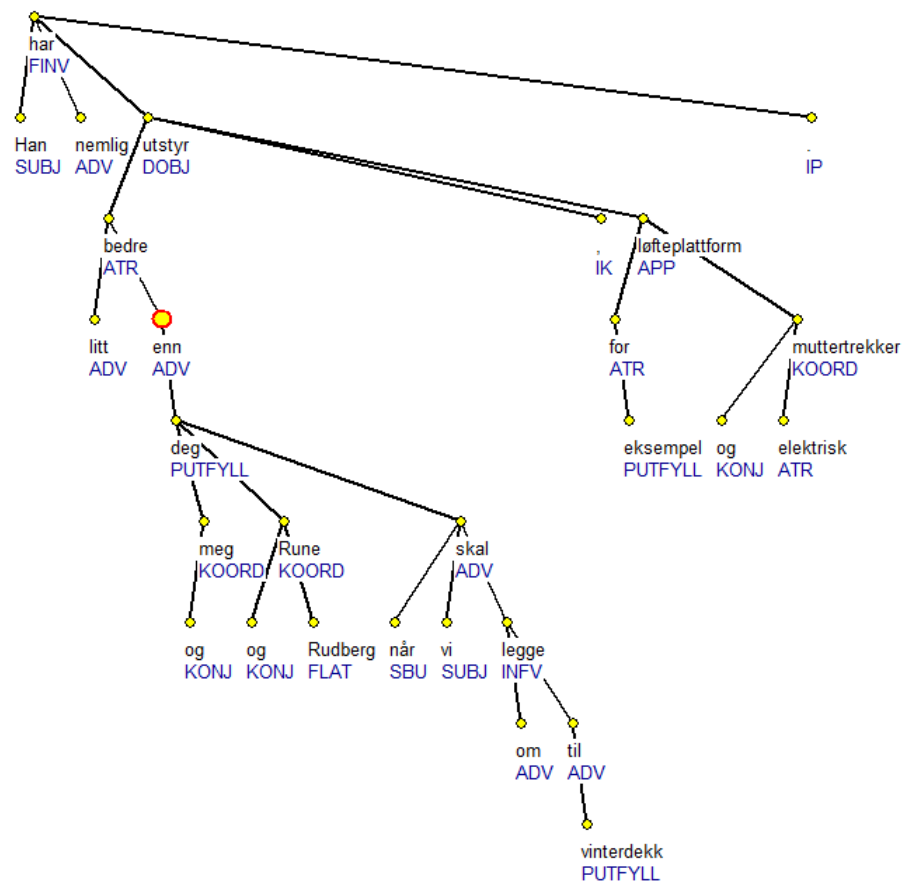
- (5.19) (a) *Dette er ikke en ny debatt, dette kommer ikke som*
 this is not a new debate, this comes not as
 julekvelden på kjerringa.
 Christmas eve+the on woman+the
 'This is not a new debate, it does not come as Christmas eve [comes]
 to the woman.' (Stortinget 001)

(b)



- (c) *Han har nemlig litt bedre utstyr enn deg og meg og Rune Rudberg når vi skal legge om til vinterdekk, for eksempel løfteplattform og elektrisk muttetrekker.*
 he has in fact somewhat better equipment than you. acc
 and me and Rune Rudberg when we shall lay about
 to winter tyres for example elevating platform and electric
 nut runner
 'He has in fact somewhat better equipment than you and me and
 Rune Rudberg when we change to winter tyres, e.g. an elevating
 platform and an electric nut runner.' (Dagbladet 004)

(d)



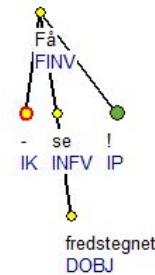
Note that the forms *deg* and *meg* in (5.19 c), despite the fact that they morphologically speaking are object forms, act as subjects in the elliptical clause.

5.2 Punctuation

5.2.1 Sentence punctuation

With sentence punctuation, we mean punctuation that ends a complete sentence, such as period, exclamation mark, question mark, etc. These will be dependents on the top node of the tree and have the function IP:

- (5.20) (a) - *Få se fredstegnet!*
 get.imp see peace sign+the
 'Let us see the peace sign!' (Aftenposten 001)
- (b)



Occasionally, sentence punctuation occurs inside a sentence, e.g. in connection with quotes, as in (5.21).

- (5.21) *“Ikke gjør det!”*, sa han.
 not do it said he
 ‘“Don’t do it!”

They should be treated as sentence-internal punctuation.

Quotation marks and closing parentheses can sometimes occur after sentence punctuation, as in (5.22 a) and (5.22 b). Although they are at the end of a complete sentence, such punctuation is considered as sentence-internal punctuation.

- (5.22) (a) *“Jeg heter Per.”*
 I am-named Per
 “My name is Per.”
 (b) *(Dette tar vi ikke stilling til.)*
 this take we not position to
 ‘(We haven’t taken a stand on this.’

5.2.2 Sentence-internal punctuation

With sentence-internal punctuation we mean all punctuation which is not used to separate main clauses, such as slashes, commas, quotation marks, parentheses, etc.

Sentence-internal punctuation shall have the function IK. It has been difficult to develop precise, general rules for the analysis of the sentence-internal punctuation, as they are often used quite randomly. However, we have formulated the following two principles:

- (5.23) (a) The punctuation mark is dependent on the head of the syntactic unit of which it marks the boundary. The punctuation mark is dependent on the left or right unit, depending on the type of punctuation mark.
 (b) In cases where there are multiple solutions to the principle in 5.23 a, the punctuation mark is placed as high as possible.

In the above principles, we refer to syntactic units, not phrases, because punctuation marks often mark the boundaries of syntactic units which are not, strictly speaking, phrases. In the following, we will show how these principles can be applied.

Commas should always be dependent on a head to the left.

(5.24) *Det finnes fattige i alle land, men det er særlig*
it exists poor in all countries but it is especially
fattige mennesker i fattige land FN konsentrerer seg
poor people in poor countries UN concentrate refl
om.
about

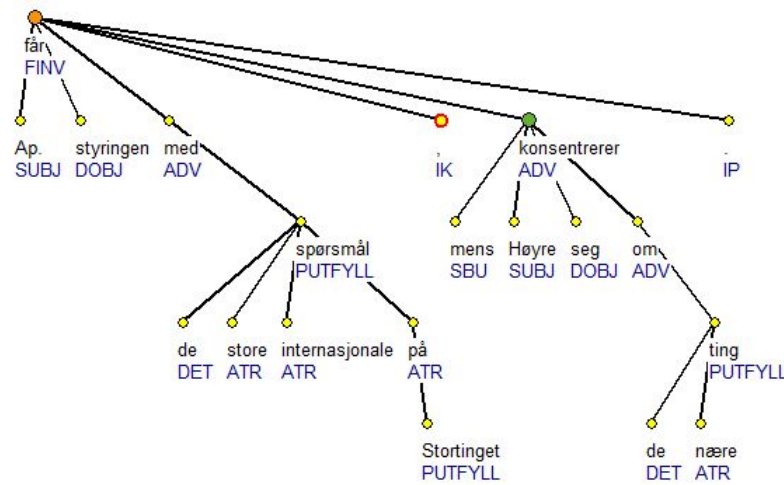
In accordance with the principle 5.23 a, the comma should be dependent on the head of the unit of which it marks the boundary. In (5.24), the comma separates two conjuncts. Since a comma always should be dependent on a head to the left, it is in this case dependent on the first conjunct:

[illegible]

(5.26) (a) *Da Aftenposten forsøkte å få ham til å utdype*
 when Aftenposten tried to get him to to elaborate
muligheten, svarte Hagen bare:
 possibility+the answered Hagen just
 'When Aftenposten tried to get him to elaborate on the possibility,
 Hagen just answered:' (Aftenposten 001)

- In (5.26 a) the comma marks the boundary of an adverbial clause at the beginning of the sentence, and it is the adverbial clause which takes the comma as dependent. In (5.26 b), where the clause is at the end of the sentence, it is the matrix clause's verb which is head for the comma:

(b)



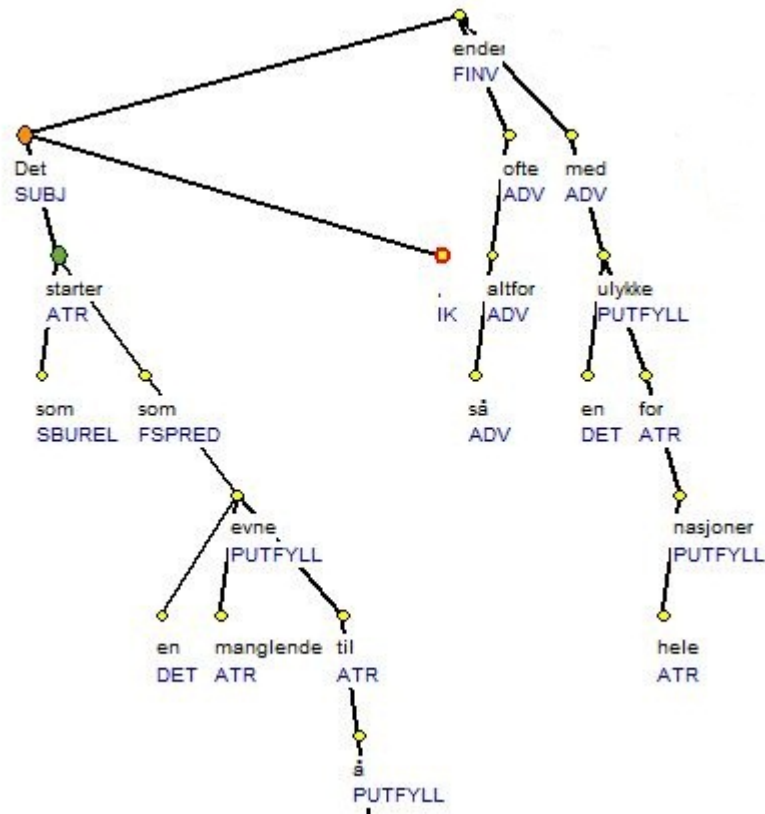
Commas often mark the boundary of an element early in the sentence which is heavy, either because it contains many words or because it contains a finite verb. In (5.28), the subject contains a relative clause, and the author has therefore put in a comma.

- (5.28) *Det som starter som en manglende evne til å*
 that which starts as a lacking ability to to
opprettholde ett menneskelivs verdighet, ender så alt for ofte
 sustain one human life's dignity ends so all for often
med en ulykke for hele nasjonen [...].
 with an accident for whole nation+the

'That which starts as a lacking ability to sustain the dignity of one persons life, ends all too often with an accident for the whole nation.'
 (Aftenposten 001)

In this case, principle 5.23 a gives several possible solutions: Either the punctuation mark must be dependent on the head of the relative clause or on the matrix subject. Here, we use principle 5.23 b and put the comma on the subject:

- (5.29)



A semicolon or a sentence-internal colon, question mark or a period shall have the same analysis as commas.

5.2.4 Quotation marks and parentheses

Quotation marks and parentheses consist of two corresponding elements on each side of a syntactic element, see (5.30 a) and (5.30 b).

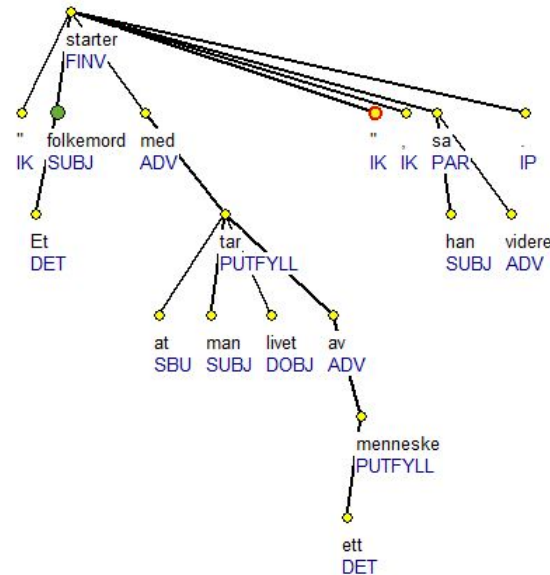
- (5.30) (a) *“Et folkemord starter med at man tar livet av
a genocide starts with that one takes life+the of
ett menneske”, sa han videre.
one person said he further
“A genocide starts with the murder of one person”, he then said.’*
(Aftenposten 001)

- (b) *Han nevner ikke konkurransen med oljebjørnen i
he mentions not competition+the with oil bear+the in
øst, den nystartede KHL (Kontinental Hockey
east the recently-established KHL Kontinental Hockey
League) [...].
League*

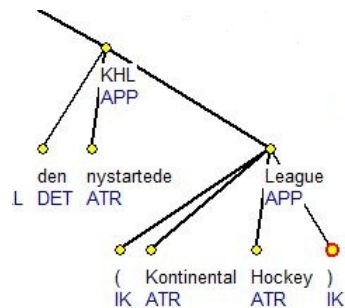
‘He doesn’t mention the competition with the oil bear in the east, the recently established (Kontinental Hockey League).’ (Dagbladet 001)

Both quotation marks or both parentheses will be dependent on the head of the syntactic element they mark the boundary of. The analysis of (5.30 a) and (5.30 b) is therefore as follows:

(5.31) (a)



(b)



5.3 Parenthetical expressions

Some constructions are difficult to analyze as having a specific syntactic function in the sentence. In such cases, we use the function PAR (parenthetic expression). Parentheses around a word or phrase is neither a necessary nor sufficient condition for the PAR function: On the one hand, parentheses are often put around phrases which have a normal syntactic function. The phrase in parentheses in (5.32 a) should be ATR on *Neumann*, in (5.32 b), it should be ATR on *Hansen* (cf. 4.1.2):

- (5.32) (a) *Ein som har skrive godt om dette, er den alt for*
 one who has written well about this is the all too
tidleg avlidne Franz Neumann (død i 1954, 54 år
 early deceased Franz Neumann dead in 1954 54 years
gamal).
 old
 'Someone who has written well about this, is the all too early
 deceased Franz Neumann (died in 1954, 54 years old).' (Dag og Tid
 001)
- (b) *Paul Hansen (27) kommer fra Haugesundsbandet*
 Paul Hansen 27 comes from Haugesund band+the
Cortina (...).
 Cortina
 'Paul Hansen (27) comes from the Haugesund band Cortina.'
 (Dagbladet 003)

On the other hand, there are phrases which do not appear with parentheses around them, but still should be PAR, as we shall see in the following.

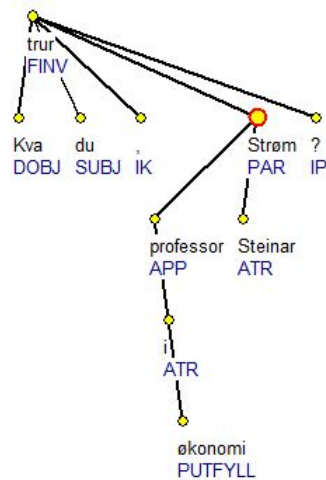
If we use the PAR function on a word, it means that the word does not have a clear syntactic function in the sentence, or that we have chosen not to take a position on what the syntactic function is. Since PAR-words have not received a truly syntactic function, they should not be dependent on any other word than the head of the sentence. Put another way, if there are reasons to make a word dependent on another word than the head of the sentence, it should not have the function PAR.

5.3.1 Address

- (5.33) *Kva trur du, professor i økonomi Steinar Strøm?*
 what think you professor in economics Steinar Strøm
 'What do you think, professor in economics Steinar Strøm?' (Dag og Tid
 002)

Names, titles and the like used in addresses, as in (5.33), has a discursive or pragmatic function, but does not fill any clear syntactic function. Therefore, the function PAR is used:

- (5.34)



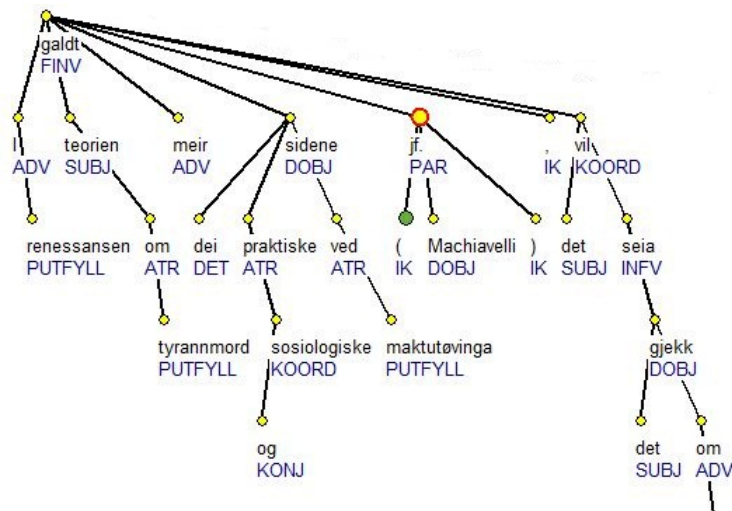
5.3.2 Parenthetical clauses

Sentences with the finite verb in parentheses are sometimes either at the end of another sentence or inserted in it, as in (5.35):

- (5.35) *I renessansen galdt teorien om tyrannmord*
 in Renaissance+the concerned theory+the about tyrant murder
meir dei praktiske og sosiologiske sidene ved
 more the practical and sociological sides+the by
maktutøvinga (jf. Machiavelli), det vil
 execution of power+the compare-with.imp Machiavelli that will
seia det gjekk om metodar som kunne koma til bruk for
 say it went about methods which could come to use for
mest mogeleg rasjonelt å nå eit mål.
 most possible rationally to reach a goal
 'In the Renaissance, the theory of tyrant murder concerned more the
 practical and sociological sides of the execution of power (see
 Machiavelli), that is, it was about methods which could be used to reach
 a goal in the most rational way.' (Dag og Tid 001)

It is often not easy to see that such clauses are subordinate to any of the other words in the sentence. They may also have verbs in the imperative, as in (5.35), which is usually not possible in subordinate clauses, and they have main clause word order. Therefore, they shall be PAR in NDT, as the analysis in (5.36) shows:

(5.36)



5.3.3 Utterance verbs in direct speech

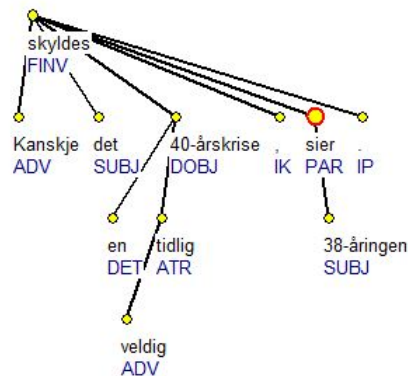
Perhaps the most common constructions with PAR in NDT are utterance verbs in direct speech, as in (5.37):

- (5.37) *Kanskje skyldes det en veldig tidlig 40-årskrise, sier*
 maybe cause.pass it a very early 40-years crisis says
38-åringen.
 38 year old+the
 'Maybe it is caused by a very early 40 years crisis, says the 38 year old man.' (Dagbladet 002)

Such sentences consist of two parts: a first part, usually a finite clause, which renders the direct speech, and another part with a finite utterance verb and its subject. It is possible to argue that the first clause is dependent on the utterance verb. Firstly, it stands in the left periphery of the utterance verb's clause. Secondly, utterance verbs are a type of verb that usually take clausal objects. However, the first sentence has main clause word order and use of indexicals is different than what is found in normal complements clauses to utterance verbs. Such a sentence is also often in a sequence of direct speech, in which the other sentences do not have utterance verbs, and the utterance verb and its subject can be omitted without rendering the sentence ungrammatical in any way.

In NDT, we have chosen to let the finite verb in the direct speech be the head, with the utterance verb as a PAR-dependent, as the analysis in (5.38) shows. This choice has primarily been taken for practical reasons: If we had let the utterance verb be the head, it had been difficult, perhaps impossible, to distinguish between these sentences and cases of indirect speech, where the same utterance verbs take an object clause. We thought it was important to distinguish these structures from each other. Also, the analysis of sentences like (5.37) would be different from other, surrounding sentences in direct speech, a consequence we thought was unfortunate.

(5.38)



5.3.4 Parenthetic verbs in the imperative

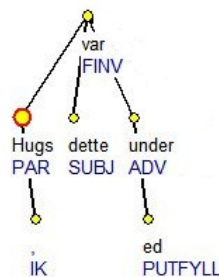
Sometimes, we find an imperative verb at the beginning of the sentence, often a thought verb such as *husk*, 'remember', *tenk*, 'think' etc...:

- (5.39) *Hugs, dette var under ed...*
 remember this was under oath

'Remember, this was under oath...' (Klassekampen-nn 001)

This construction has something in common with the parenthetic utterance verbs mentioned above: The rest of the sentence somehow seems to be an argument of the verb in the imperative. However, the verb in the imperative can be omitted without changing the meaning, and there is main clause word order in the rest of the sentence. We have therefore chosen to treat these imperatives the same way as parenthetical utterance verbs: The verb in the imperative becomes PAR:

(5.40)



5.3.5 Utterance verbs at the beginning of the sentence

Often, an utterance verb is followed by a colon and then a main clause:

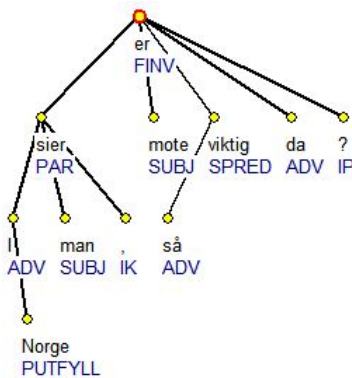
- (5.41) *Naturkraft sier selv i en pressemelding: "Vi håper å kunne skaffe en del av de CO2-kvotene vi trenger [...]".*
 Naturkraft says themselves in a press release we hope to
 could obtain a part of the CO2 quotas we need

'Naturkraft says in a press release: "We hope to be able to obtain a part of the CO2 quotas we need."' (Stortinget 002)

In such cases, there is a sentence boundary after the colon.

Occasionally, however, a comma is used instead of a colon, as in (5.42 a). Then we cannot put in a sentence boundary, since a comma is not a sentence-separating punctuation mark. In such cases, the utterance verb gets the PAR function, as the analysis in (5.42 b) shows:

- (5.42) (a) *I Norge sier man, er mote så viktig da?*
 In Norway says one is fashion so important then
 'In Norway we say, is fashion really so important?' (Dagbladet 014)
- (b)



We only use this analysis, however, if it is absolutely clear from the word order etc., that what follows after the comma is a main clause and cannot be a subordinate clause. In (5.42 a), we have for example question word order, which is not possible in subordinate clauses. If we are in doubt, we always choose a subordinate clause analysis.

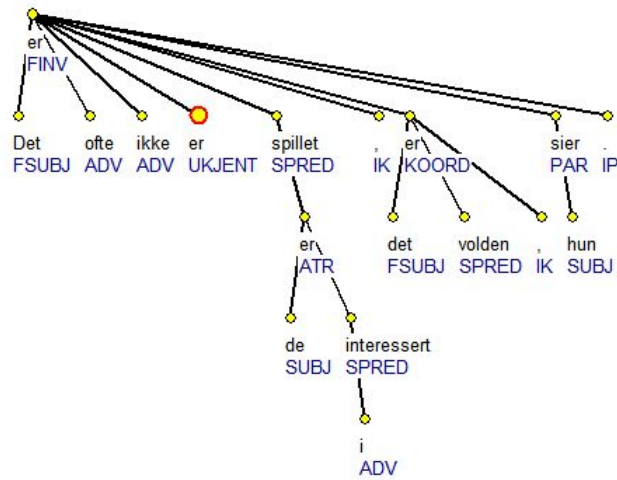
5.4 Superfluous words in the sentence - UKJENT

Occasionally, superfluous words occur in the sentence, which are there because of errors, etc., as the second instance of *er* in (5.43).

- (5.43) *Det er ofte ikke er spillet de er interessert i, det*
 it is often not is game+the they are interested in it
 er volden, sier hun.
 is violence+the says she
 'It is often not is the game they are interested in, it is the violence, she says.' (Klassekampen 005)

Since such words are ungrammatical, we wanted to give them a function that is not used in other instances. We have called this function UKJENT (*unknown*). UKJENT-dependents should always have the main verb as head. We have tried to ensure that this function is used only when no other analysis is possible. Users who do not want such ungrammatical sentences in their material, can sort out all sentences containing UKJENT-dependents.

(5.44)



5.5 *Siden*

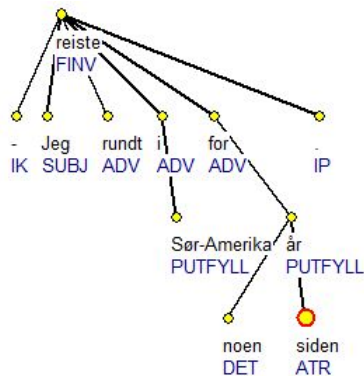
5.5.1 *Siden* as an adverb

The adverb *siden*, 'since', is often found after adverbial nouns, as in (5.45). Often (but not always), these nouns are complements to the preposition *for*, 'for'.

- (5.45) -Jeg reiste rundt i Sør-Amerika for noen år siden.
 I traveled around in South America for some years since
 'I traveled around in South America, some years ago.' (Dagbladet 003)

It is not obvious how *siden* should be analyzed in such cases. We have chosen to let *siden* be attribute to the noun in such cases, as shown in (5.46).

(5.46)



5.5.2 *Siden* as a preposition

Siden can also be a preposition, as in (5.47):

- (5.47) *Etter å ha jobbet seks og syv dager i uka*
 after to have worked six and seven days in week+the
siden 1992, har jeg endelig den friheten og kontrollen
 since 1992 have I finally the freedom+the and control+the
jeg har drømt om.
 I have dreamt about
 'After having worked six and seven days a week since 1992, I have finally
 the freedom and control I have dreamt of.' (Dagbladet 001)

The morphological tagger does not allow preposition as a part of speech for *siden*. This must therefore be corrected manually, cf. 1.2.2.

5.5.3 *Siden* as a complementizer

Siden can also be a complementizer, as in (5.48).

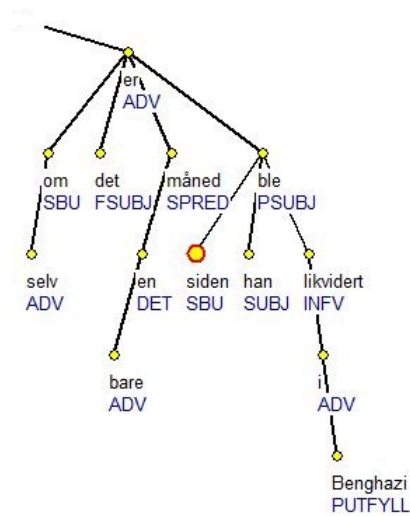
- (5.48) *Siden jeg spilte den melodien ofte for andre, må jeg*
 since I played the melody+the often for others must I
ha vært ganske trygg på den.
 have been quite confident on it
 'I must have been quite confident about that tune, since I played it often
 for others.' (Dagbladet 003)

A subordinate clause introduced by *siden* occurs in constructions such as (5.49), with formal subject, copular verb and predicative.

- (5.49) *Selv om det bare er en måned siden han ble likvidert*
 even if it just is a month since he became liquidated
i Benghazi [...].
 in Benghazi
 'Even if it is just a month since he was liquidated in Benghazi ...'
 (Klassekampen 004)

Such constructions are very similar to clause-anticipating constructions, which we have described in 3.2.3 (cf. Faarlund et al. 1997, 1071). One difference from the usual cases of clause-anticipating constructions is that the subordinate clause in examples like (5.49) cannot be made a subject. We have still chosen to analyze such sentences as clause-anticipating constructions. The clause should, in other words, have the function PSUBJ:

- (5.50)



Subordinate clauses with the complementizers *før* and *til* can behave the same way, cf. Faarlund et al. 1997, 1071.

5.6 Quantity indications of the type *to og et halvt år*

See (5.14 a) and (5.14 b) in subsection 5.1.3.1.

5.7 *Ha vanskelig for*

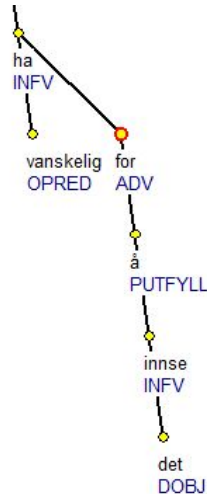
How should we analyze constructions such as (5.51), with *ha*, 'have', an adjective and a prepositional phrase with *for*, 'for'?

- (5.51) *Selv om mange i Kr.F og Venstre ser ut til å ha vanskelig for å innse det, er det et faktum at Fr.P. blir det viktigste støttepartiet for Bondeviks andre regjering.*
 even if many in Kr.F and Venstre looks out to to have
 difficult for to realize it is it a fact that Fr.P.
blir det viktigste støttepartiet for Bondeviks andre regjering.
 becomes the important.sup support party for Bondevik's second
 government
 'Even if many in Kr.F and Venstre seem to have difficulties realizing it,
 it is a fact that Fr.P. will become the most important support party for
 Bondevik's second government.' (Aftenposten 001)

Ha generally takes a direct object, and it can apparently be a possibility to let the adjective receive this function. But in that case we expect an implied noun which the adjective modifies, and we do not think this is the correct analysis here. We believe it is more appropriate to assume that *ha* has the prepositional phrase as a sort of prepositional object. Since we do not distinguish between compulsory and non-compulsory prepositional phrases in our analysis, *for* gets

the function ADV. The adjective, *vanskelig* in our example, we analyze as an object predicative:

(5.52)



5.8 Comparative constructions

Here we will treat the analysis of various types of comparative constructions. In 5.8.1, we provide the first general rule for what should be the head of comparative constructions. In 5.8.2 and 5.8.3, we explain our analysis of constructions such as *slik/samme som* and *jo mer, jo bedre*.

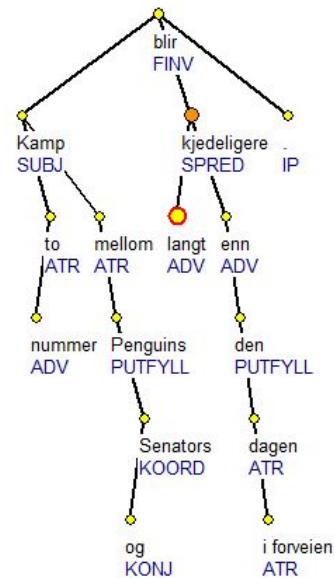
5.8.1 The head in comparative constructions

The general rule for the analysis of comparative constructions is that the word indicating the comparison is the head of what we might call *the comparison phrase*. Adjectives in the comparative is perhaps the most typical example. Below, we see that *kjedeligere*, 'boring.comp', is the head of the prepositional phrase introduced by *enn*, 'than' (i.e. the comparison phrase). We also see that the comparison phrase should have the function ADV (see 3.5 for a description of adverbials in general).

- (5.53) (a) *Kamp nummer to mellom Penguins og Senators*
 match number two between Penguins and Senators
blir langt kjedeligere enn den dagen i
 becomes much boring.comp than that day+the before
forveien.

'Match number two between Penguins and Senators becomes much more boring than the one the day before.' (Dagbladet 001)

(b)



The general rule we have presented here is in line with the principles of Dependency Grammar because the comparison phrase has no place in the sentence without the adjective in the comparative (or other phrases with a similar function, as we shall see further below). The comparison phrase depends on its head, but the reverse does not apply. This is why (5.54 a) below, where the head is omitted, is ungrammatical, while (5.54 b), where the head remains, whereas the comparative clause is omitted, is in good.

- (5.54) (a) **Kamp nummer to blir enn den dagen i*
 match number two becomes than that day+the before
forveien

'*Match number two becomes than the one the day before.'

- (b) *Kamp nummer to blir vanskeligere*
 match number two becomes difficult.comp

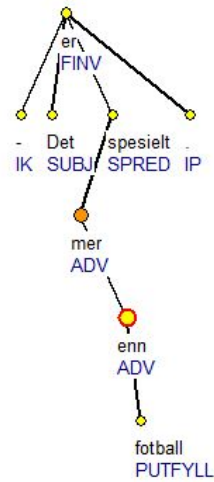
'Match number two becomes more difficult.'

A comparative adjective that very often is head for a comparison phrase is *mer*, 'more'. *Mer* can either stand alone or as ADV to another adjective (periphrastic comparative), but the analysis is the same in both cases. The following examples illustrate this:

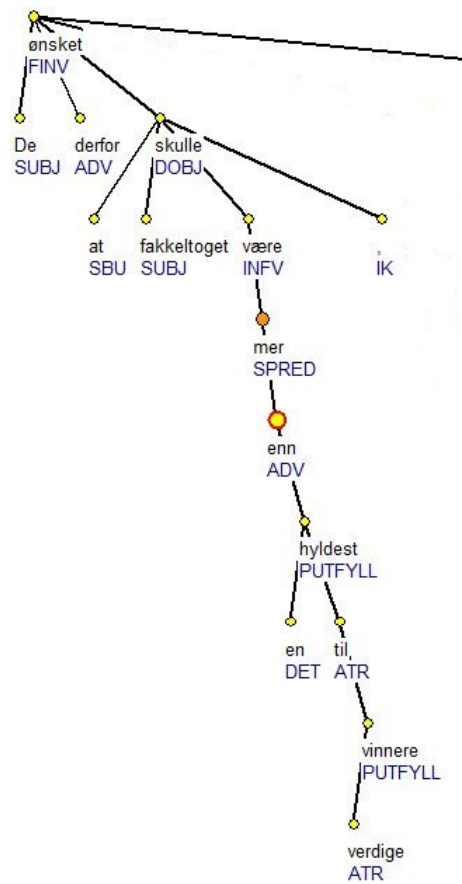
- (5.55) (a) -*Det er mer spesielt enn fotball.*
 it is more special than football

'It is more special than football.' (Dagbladet 003)

- (b)



- (c) *De ønsket derfor at fakkeltoget skulle være mer enn en hyldest til verdige vinnere [...].*
 They wished therefore that torchlight procession+the should be more than a tribute to worthy winners
 'They therefore wished that the torchlight procession should be more than a tribute to worthy winners.' (Dagbladet 001)
- (d)

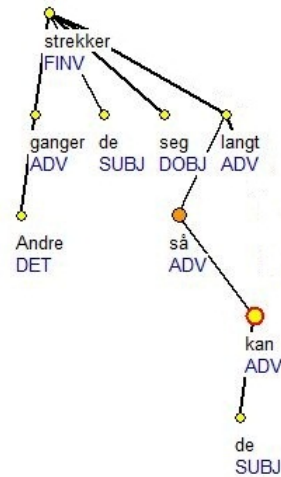


Mer is the head in the comparison construction in both (5.55 a), where it modifies *spesielt* as ADV, and (5.55 c) where it is DOBJ.

Comparative adjectives are not the only ones which can be head in comparison constructions. We would particularly like to mention two other words that often has this role, namely the adverbs *like*, 'as' and *så*, 'so'. *Like* and *så* modify adjectives and introduces comparison in sentences such as (5.56 a) - (5.56 g) below. The figures show how *så* and *like* function as heads of comparison phrases.

- (5.56) (a) *Andre ganger strekker de seg så langt de kan [...].*
 other times stretches they refl so long they can
 'Other times, they went out of their way as much as they could.'
 (Aftenposten 001)

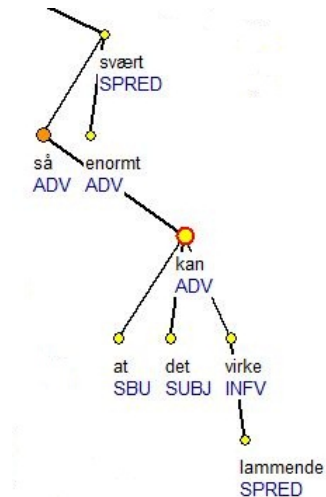
(b)



- (c) *For noen er det å komme gjennom den tilgivelsesbiten så enormt svært at det kan virke lammende.*
 for some is that to come through the forgiveness part+the so hugely huge that it can function paralyzing

'For some, coming through the forgiveness part is so hugely difficult that it can be paralyzing.' (Dagbladet 001)

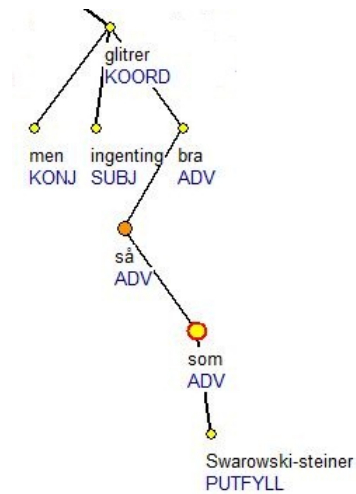
(d)



- (e) *[...] men ingenting glitrer så bra som Swarovski-steiner.*
 but nothing sparkles so good as Swarovski stones

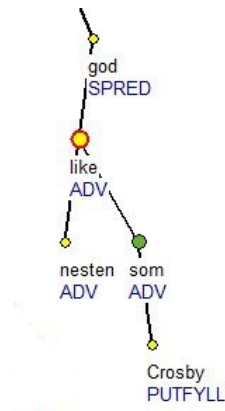
'But nothing sparkles as nicely as Swarovski stones.' (Dagbladet 002)

(f)



- (g) [...] 22-åringen som ifølge dem som kan det
 22 year old+the who according to them who can it
 er nesten like god som Crosby [...].
 is almost as good as Crosby
 'the 22 year old man who according to those who know, is almost as
 good as Crosby' (Dagbladet 001)

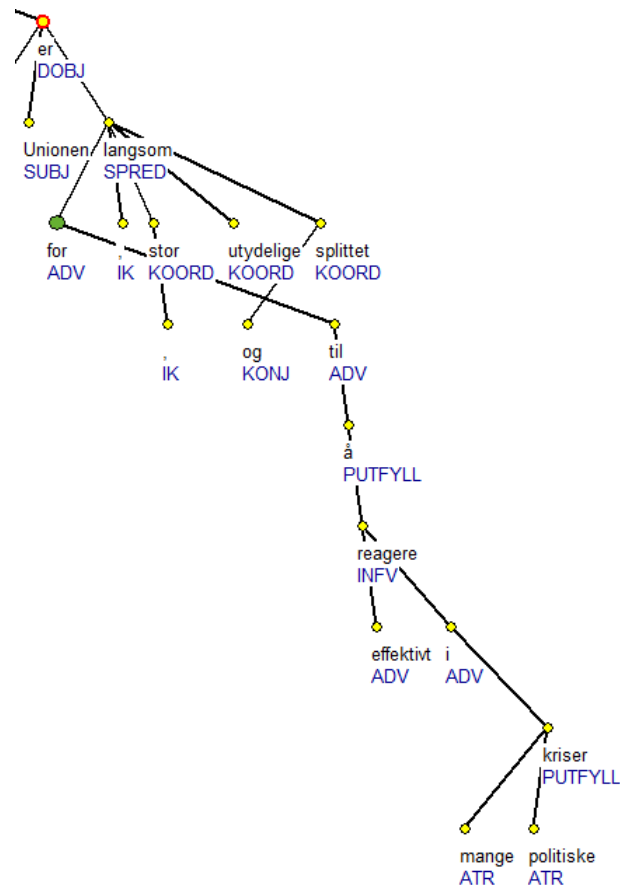
(h)



Comparative constructions with *for...til* and *nok...til* are treated the same way:

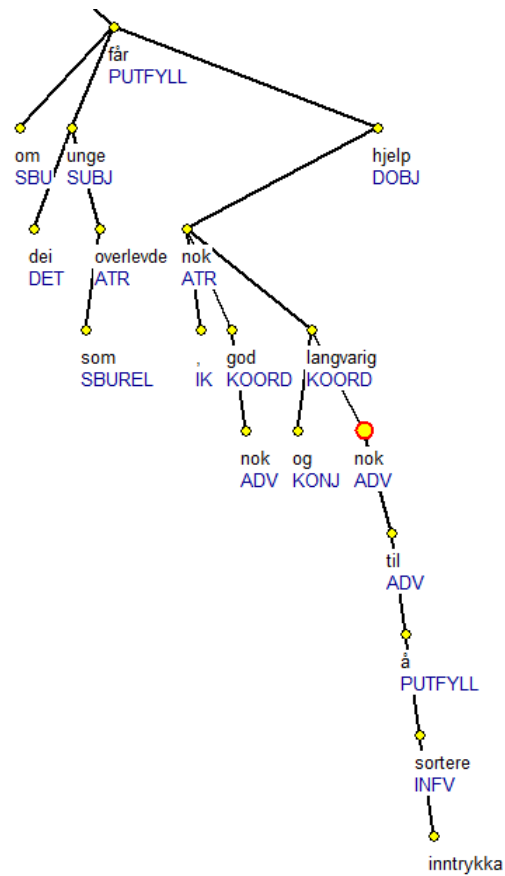
- (5.57) (a) *Unionen er for langsom, stor, utydelige og splittet til
 union+the is too slow big vague and split to
 å kunne reagere effektivt i mange politiske kriser.
 to could.inf react effectively in many political crises
 'The union is too slow, big, vague and split to react effectively in
 many political crises.'* (Aftenposten 009)

(b)



- (c) [...] *om dei unge som overlevde får nok, god nok*
 if the young who survived get enough good enough
og langvarig nok hjelp til å sortere inntrykka.
 and lengthy enough help to to sort impressions
 '... if the youths who survived get enough help, good enough help
 and and help which is lengthy enough to sort out the impressions.'
 (Klassekampen-nn 003)

(d)

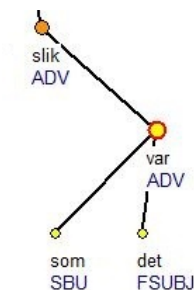


5.8.2 *Slik/samme som*

We assume that some adjectives can be heads for comparison phrases without being in the comparative. A common example of this is *slik* followed by a comparison phrase introduced by *som*:

- (5.58) (a) *Vi kunne ikke ha fortsatt slik som det var.*
 we could not have continued such as it was
 'We could not have continued as it was.' (Dagbladet 003)

(b)



Slik is often followed by *at* in consequence and purpose clauses, as in (5.59):

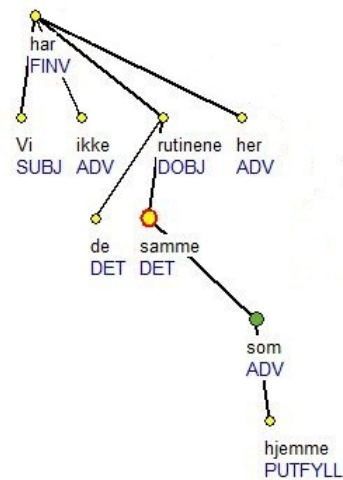
- (5.59) *Dermed vil strålene som treffer metallplata i*
 consequently will rays+the which meet metal plate+the in
veska sende tilbake et omriss av plata og bokstavene,
 bag+the send back a contour of plate+the and letters+the
slik at beskjeden kommer fram.
 such that message+the comes forward
 'Consequently, the rays which meets the metal plate, will send back a
 contour of the plate and letters, so that the message is spread.'
 (Dagbladet 001)

Note that sentences like (5.59) are not comparative constructions, and thus should not have the same analysis as *slik som*. We consider *slik at* a kind of composite complementizer and let *slik* be ADV-dependent on *at*. See 3.10.2 for more on dependents to complementizers.

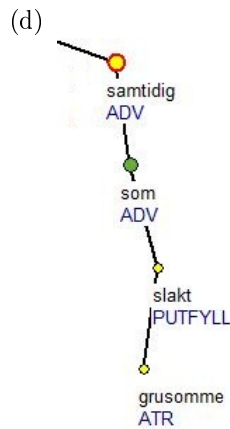
Some other words that typically can head comparison phrases, are the determiner *samme*, 'same', and the adjective *samtidig*, 'simultaneous'. We see examples of this below.

- (5.60) (a) *Vi har ikke de samme rutinene her som hjemme [...].*
 we have not the same routines here as at home
 'We do not have the same routines here as at home.' (Dagbladet 001)

(b)



- (c) *Få, om noen norske regissører, har fått så mange*
 few if any Norwegian directors have got so many
strålende filmanmeldelse - samtidig som grusomme
 brilliant film reviews simultaneously as terrible
 slakt.
 slamming
 'Few, if any Norwegian directors have received so many brilliant film
 reviews - at the same time as terrible slamming.' (Dagbladet 001)



5.8.3 *Jo mer, jo bedre*

Occasionally comparative constructions with two comparative adjectives occur, both of which are modified by *jo* or *dess/di*, as in (5.61 a) and (5.61 b). In such cases, it is not clear what is the head, and what is the dependent.

(5.61) (a) -*Jo mer, jo bedre.* (*Dagbladet* 002)

jo more jo better

'The more the better.'

(b) -*Vi forventer innflytelse på statsbudsjettet og vi har gjort det klart at jo større vi blir jo større innflytelse skal vi ha, sier Schmidt-Nielsen.*

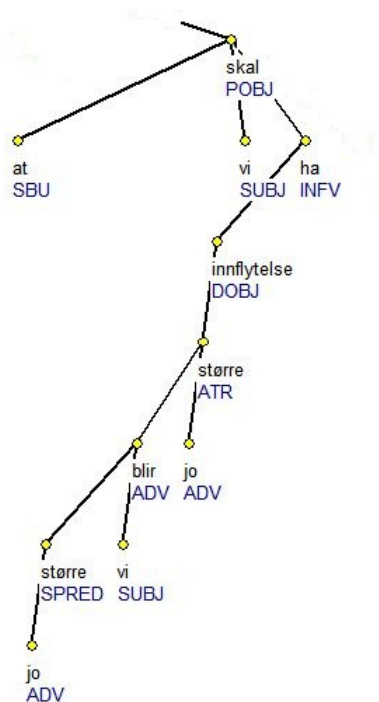
we expect influence on state budget+the and we have made it clear that jo bigger we become jo greater influence shall we have says Schmidt-Nielsen

'We expect influence on the state budget, and we have made it clear that the bigger we become, the greater influence we shall have.'

(*Klassekampen* 003)

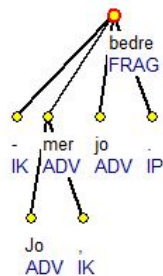
The word order in (5.61 b) gives an idea of how this type of comparison should be analyzed: While the first part of the comparison, *jo større vi blir*, has a subordinate clause word order, the second part, *jo større innflytelse skal vi ha*, has main clause word order (see also Faarlund et al. 1997, 1087-1088). Therefore, there are probably reasons for making the second comparative adjective the head. The analysis is thus as follows:

(5.62)



We have chosen to analyze examples like (5.61 a), where there is no verb, the same way:

(5.63)



5.9 Mixed quotation

It is quite common that quotations occur inside non-quoted sentences. The quote can be in the same language as the text in general, as in (5.64 a), or in another language, as in (5.64 b). This is called *mixed quotation* in linguistic terminology (cf. Maier (2007)).

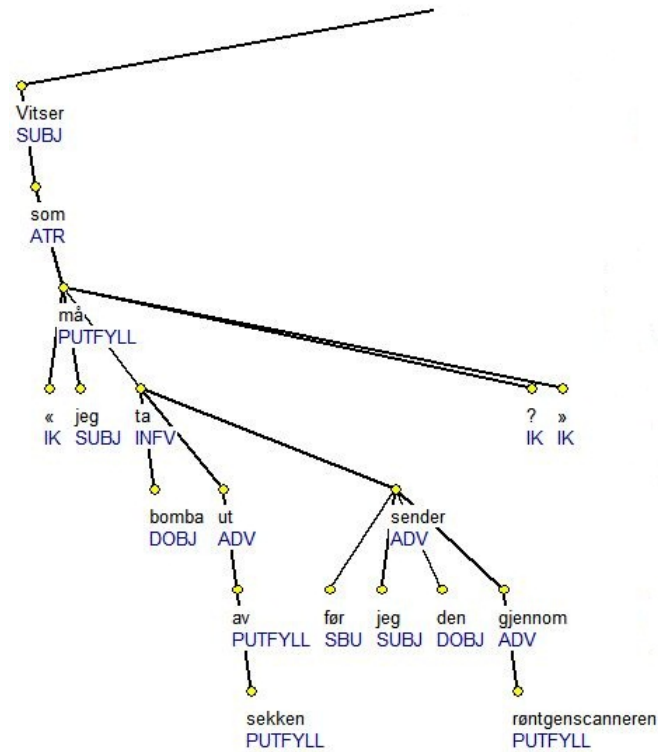
- (5.64) (a) *Vitser som «må jeg ta bomba ut av sekken
jokes as must I take bomb+the out of backpack+the
før jeg sender den gjennom røntgenscanneren?» har
before I send it through x-ray machine+the have*

vært populære blant skjemtegaukene i security-køen,
 been popular among joking cuckoos+the in security line+the
om enn ikke så populære blant bryske sikkerhetsvakter.
 if than not so popular among brusque security guards.
 'Jokes such as "Do I need to take the bomb out of the backpack
 before I send it through the x-ray machine?" have been popular
 among the jokers in the security line, if not so popular among brusque
 security guards.' (Dagbladet 001)

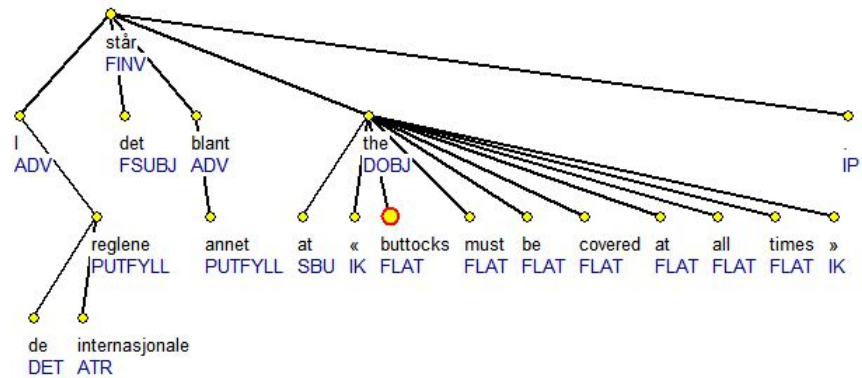
- (b) *I de internasjonale reglene står det blant annet at*
 in the international rules stands it among other that
"the buttocks must be covered at all times".
 eng eng eng eng eng eng eng eng
 'In the international rules, it says, among other things, that "the
 buttocks must be covered at all times".' (Dagbladet 002)

We have no separate analysis for quotations in Norwegian and mixed quotations in Norwegian, as in (5.64 a), is no exception. These should have a standard internal structure, cf. (5.65 a). Mixed quotations in foreign languages should be tagged as *ukjent* in the morphology (cf. 1.2.3.1) and have a flat structure in the syntax (see 4.5). Note that the Norwegian words that do not belong to the quote, but which depend on it syntactically, such as *at* in (5.65 b), should be dependent on the head of the quote. The function should be the same as if it were dependent on a Norwegian word, in this case SBU.

(5.65) (a)



(b)



5.10 Garpe-genitive

Garpe-genitive (*garpegenitiv* in Norwegian) is the popular term for genitive expressions such as *Pål sine høner*, 'Pål refl-poss hens'. We analyze garpe-genitive as an extension of the analysis of conventional genitive (4.1.3.1), as well as an extension of the possessive reflexive, *sin* (4.1.1). Both regular genitives and *sin* will be dependent on the possessed noun, and have the function DET. Based on

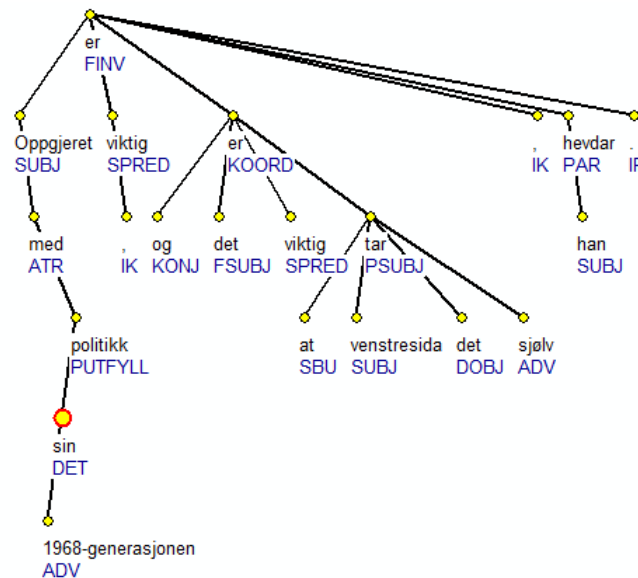
this, we will consider *Pål sine* a complex possessive phrase, which is dependent on the possessed noun, *høner*. We analyze *sin* as the head of the possessive phrase, and we give it the function DET. The possessor noun, *Pål*, becomes dependent on *sine* and gets the function ADV.

In example (5.66 a), we see how this analysis works in practice.

- (5.66) (a) *Oppgjæret med 1968-generasjonen sin politikk*
 confrontation+the with 1968 generation+the refl-poss politics
er viktig, og det er viktig at venstresida tar
 is important and it is important that left side+the takes
det sjølv, hevdar han.
 it self claims he

'The confrontation of the 1968 generation's politics is important, and it is important that the left side itself is engaged in it, he claims.'
 (Klassekampen-nn 001)

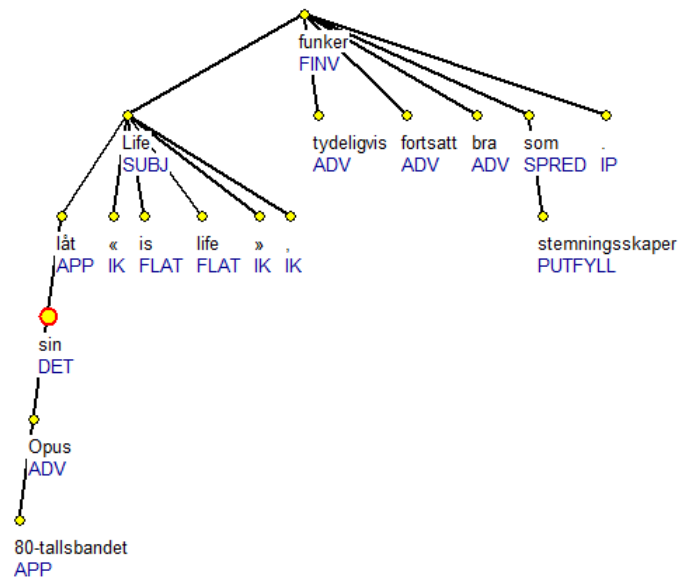
(b)



- (c) *80-tallsbandet Opus sin låt Life is Life, funker*
 80ies band+the Opus refl-poss tune Life is Life works
tydeligvis fortsatt bra som stemningsskaper.
 apparently still well as ambience creator

'The 80ies band Opus' tune Life is Life can apparently still be used to create an ambience.' (Dagbladet 001)

(d)

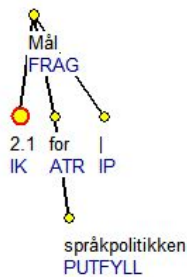


5.11 Section references etc.

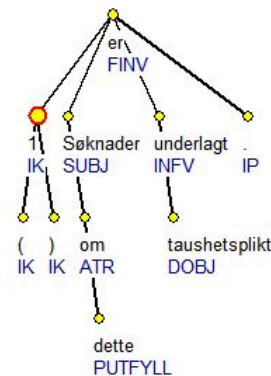
Numbers and letters at the beginning of sentences which are used as section references, in enumerations etc., get the function IK:

- (5.67) (a) *2.1 Mål for språkpolitikken/*
 2.1 goals for language policy+the
 '2.1 Goals for the language policy' (Mål og meining 002)
- (b) *(1) Søknader om dette er underlagt*
 (1) applications about this are subjected
taushetsplikt.
 duty of confidentiality
 '(1) Applications for this are subject to duty of confidentiality.'
 (Aftenposten 011)

(c)



(d)



5.12 Tough movement

Tough movement is a construction in which a nominal phrase that semantically speaking is an argument in a subordinate infinitive construction, fills the subject position in the matrix clause (cf. Faarlund et al. (1997, 1029-1032)). The following examples illustrate this phenomenon:

- (5.68) (a) *Den låge pensjonsalderen er vanskeleg å skjøna [...].*
 the low pension age+the is hard to understand
 'The low pension age is hard to understand.' (Dag og Tid 003)
- (b) *Det var grusomt å høre på [...].*
 it was awful to listen to
 'It was awful to listen to.' (Aftenposten 012)
- (c) *Resultatet er knapt til å unngå:*
 result+the is barely to to avoid
 'The result is barely possible to avoid.' (Aftenposten 009)
- (d) *Filmen kosta 25.000 dollar å laga [...].*
 film+the costed 25.000 dollar to make
 " (Klassekampen-nn 002)

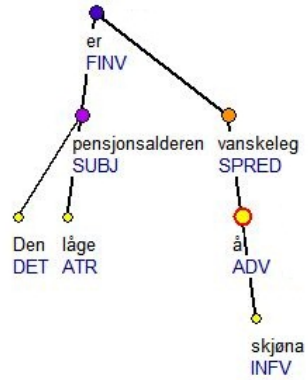
Tough movement is relatively common with a copular verb and an adjectival subject predicative, as in (5.68 a) and (5.68 b). In (5.68 a), the subject predicative *vanskelig* is followed by the infinitival clause *å skjøna*. *Den låge pensjonsalderen*, which is the subject of the matrix clause, is, semantically speaking, the object of *skjøna*. In (5.68 b), the subject predicative *grusomt* is followed by *å høre på*. *Det*, the subject of the matrix clause, is also the prepositional complement to *på*. The subject predicative does not need to be adjectival. (5.68 c) is an example of tough movement with a copular verb and a preposition as predicative.

Although tough movement is common in non-verbal predicates, it does occasionally occur also with verbal predicates, as with the verb *koste* in (5.68 d). The subject of the matrix clause, *filmen*, is also the object in the infinitival clause *å laga*.

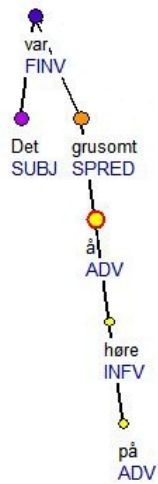
The argument that serves as the subject of the matrix clause and as an argument in the subordinate infinitival clause, will in our analysis be SUBJ-dependent on the matrix verb. In sentences like (5.68 a) and (5.68 b), we have

chosen to let the infinitive be ADV-dependent on the subject predicative. In the case of tough movement with verbal predicates, as in (5.68 d), we have chosen to let the infinitive be ADV-dependent on the matrix verb:

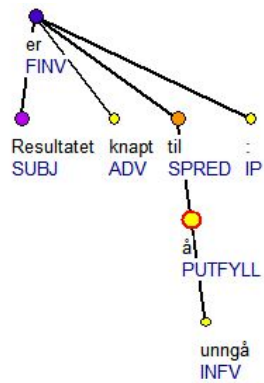
(5.69) (a)



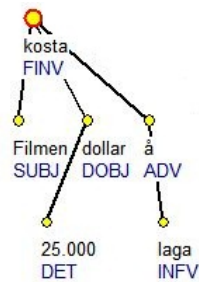
(b)



(c)



(d)



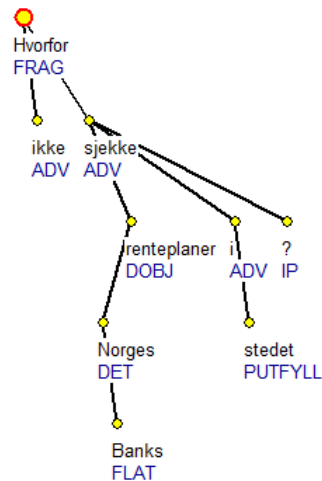
5.13 Fragmentary wh-questions

Sometimes, interrogative pronouns occur in sentences without a finite verb, eg. *hvorfor det?*, 'why that', *hvorfor ikke?*, 'why not' or *hva med en øl?*, 'what with a beer = what about a beer'. We analyze these as fragmentary sentences with the wh-word as the head. The remaining phrases in the sentence are dependent on the wh-word with the function ADV:

- (5.70) (a) *Hvorfor ikke sjekke Norges Banks renteplaner i stedet?*
 why not check Norges Bank's interest plans instead

'Why not check Norges Bank's interest plans instead?' (Aftenposten 009)

(b)



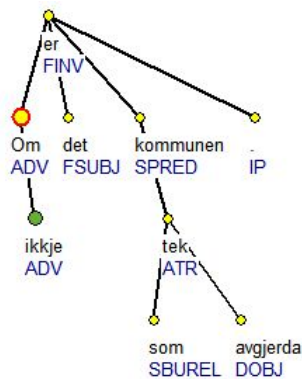
5.14 Fragmentary conditional clauses

Hvis ikke or *om ikke*, 'if not', can sometimes occur at the beginning of a sentence with the meaning 'if this doesn't happen', or something similar:

- (5.71) *Om ikkje er det kommunen som tek avgjerda.*
 if not is it municipality+the which takes decision+the
 'If not, it is the municipality which makes the decision.' (Vest-Telemark
 blad-nn 003)

We believe it is likely that this is a kind conditional clause where the verb is elided, as a complementizer usually cannot occur without a finite verb, and because it is difficult to make a paraphrase of the intended meaning without using a verb. In line with the practice with ellipses described in 5.1, we let one of the words in the elliptical construction be the head and let the second be dependent on it with the function it would have in a non-elliptical construction. In cases such as (5.71), we let the complementizer be the head and let *ikke* be dependent with the function ADV:

(5.72)



5.15 Multiple adverbials in the left periphery

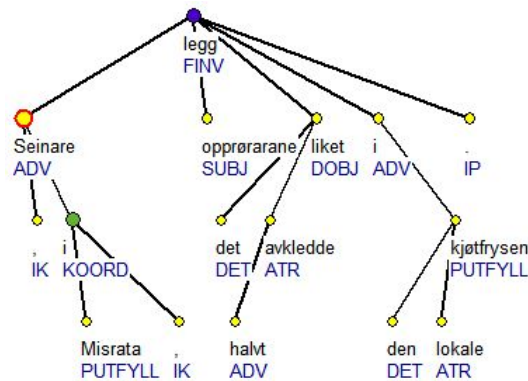
Sometimes, multiple adverbials are found in the left periphery of a matrix clause:

- (5.73) (a) *Seinare, i Mistrata, legg opprørarane det halvt avkledde*
 later in Mistrata lay rebels the half undressed
liket i den lokale kjøtfrysen.
 corps in the local meat freezer
 'Later, in Mistrata, the rebels lay the partly undressed corps in the
 local meat freezer.' (Dag og tid 001)
- (b) *Ein gong i ein by merka ho at ein mann såg*
 a time in a city noticed she that a man looked
på henne.
 at her
 'One time in a city, she noticed that a man looked at her.'
 (Vest-Telemark blad-nn 006)

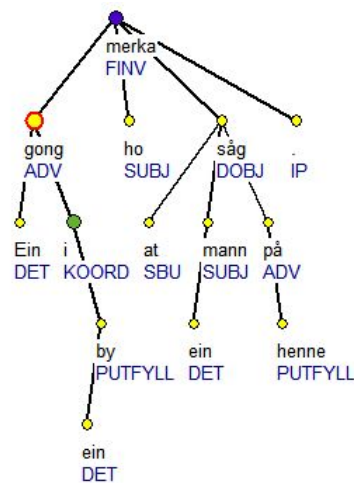
Generally, there are not multiple phrases in the left periphery of a matrix clause. However, there are several exceptions to that rule, and we could consider cases like (5.73 a) and (5.73 b) such exceptions. However, we have chosen not to

consider such examples as exceptions to the verb second rule. The reason for this is that it seems to only be adverbial phrases that can occur together in the left periphery in this way. Because of this, we think it is likely that these phrases are in a dependency relationship to each other. We have chosen to analyze them as a form of coordination:

(5.74) (a)



(b)



5.16 *Selv*

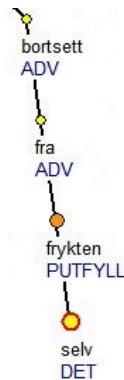
As we described in 1.2.7, *selv*, 'self', can either be tagged as an adverb or determiner in the morphology. When *selv* occurs after a nominal constituent and meets the criteria for being a determiner, it will be a DET-dependent on the head of the nominal constituent:

(5.75) (a) [...] *Vi har intet å frykte bortsett fra frykten*
 we have nothing to fear except from fear+the
selv, og vi har ingen tid å miste [...]. (Aftenposten)
 self and we have no time to loose

009)

'We have nothing to fear except fear itself, and we have no time to spare' (Aftenposten 009)

(b)

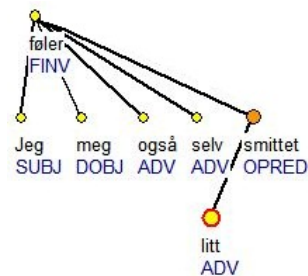


When *selv* is an adverb, it should usually be an ADV-dependent on the verb:

- (5.76) (a) *Jeg føler meg også selv litt smittet.*
 I feel myself also self somewhat infected

'I also feel a bit infected, myself.' (Aftenposten 012)

(b)

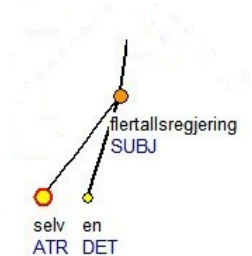
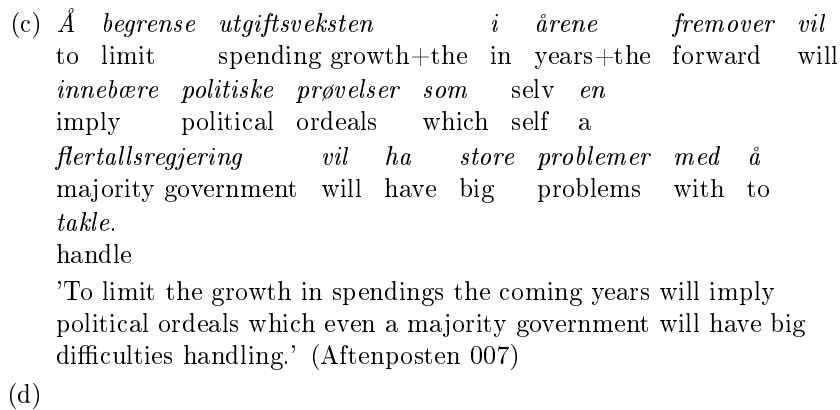


Selv with the meaning 'even' will be dependent on the word it modifies. The function should be ADV if it depends on prepositions, adverbs, etc., cf. 3.5, ATR if it depends on a pronoun or noun:

- (5.77) (a) *Men et sted går grensen, selv for Erik Poppe:*
 but a place goes limit+the even for Erik Poppe

'But there is a limit, even for Erik Poppe' (Dagbladet 001)

(b)



A subject which is modified by the quantifiers *alle*, 'all', and *begge*, 'both', and occasionally *noen*, 'some', *ingen*, 'no' and *hver*, 'each', can be separated from the quantifier, as in (5.78 a) and (5.78 b). This is often referred to as floating quantifiers (cf. Faarlund et al., 1997, 920-922).

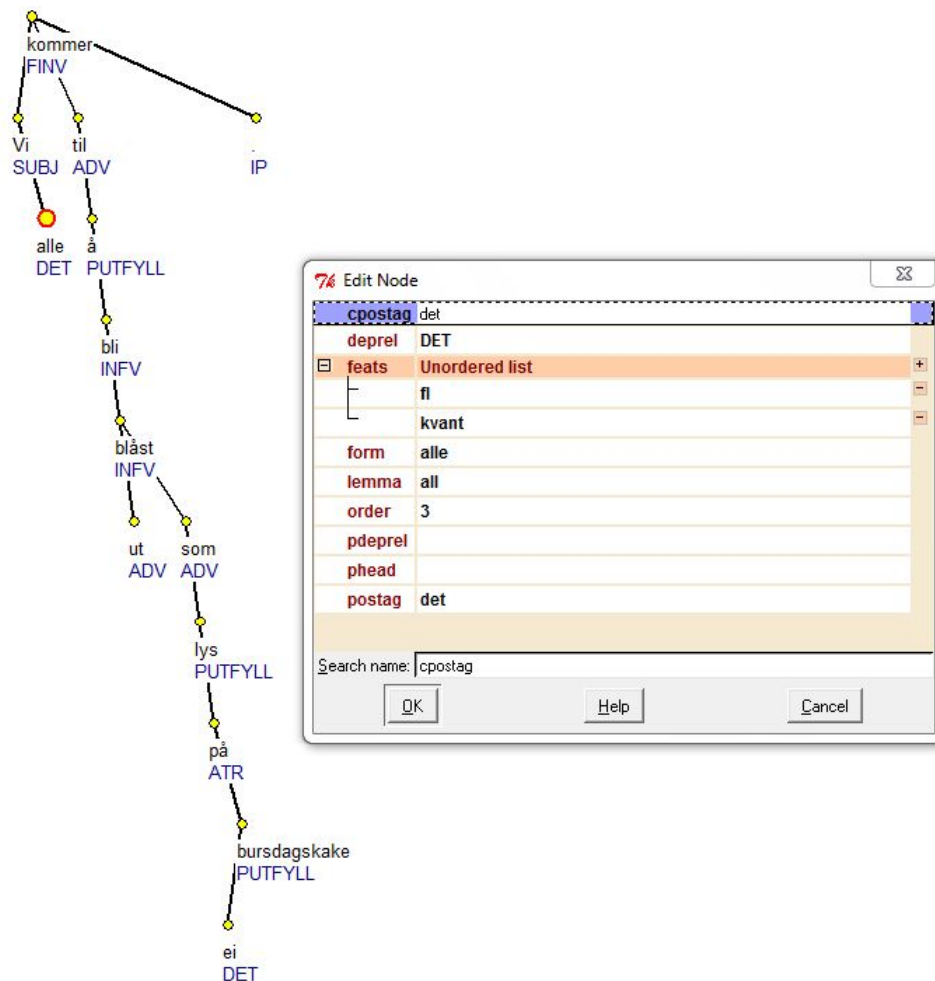
- (5.78) (a) *Vi kommer alle til å bli blåst ut som lys*
 we come all to to become blown out as candles
på ei bursdagskake.
 on a birthday cake
 'We will all be blown out, as candles on a birthday cake' (Dagbladet
 003)

- (b) *Begge måtte dei bøte med livet.*
 both must they pay with life+the

'They must both pay with their lives' (Faarlund et al., 1997, 921)

In our analysis, we have no separate analysis of quantifiers which are floating: Quantifiers should be determiners in the morphology and be dependent on the subject with the function DET, as the analysis in (5.78 a) shows:

(5.79)



Note that we do not treat *selv* as a floating quantifier, as NRG does. This choice is made of purely pragmatic reasons. For the analysis of *selv*, see 1.2.7 and 5.16.

5.18 Difficult sentences

Occasionally, you encounter difficult sentences such as (5.80). If that happens, make a cup of coffee and do the best you can.

(5.80) *At han var ein nyttig frontfigur, kan henda nyttig idiot,*
 that he was a useful frontman can happen useful idiot
men at djupna i djupøkosophien når alt
 but that depths+the in deep ecosophy+the when everything
kom til alt, var grunnare eller uklårare enn
 comes to everything was shallow.comp or muddy.comp than
ein skulle venta seg av ein som åtvara mot at når
 one should expect refl of one who warned against that when
ein ikkje såg botnen i eit vatn, kunne det koma av
 one not saw bottom+the in a lake could it come of
at vatnet var uklårt.
 that lake+the was muddy

'That he was a useful frontman, maybe a useful idiot, but that the depths of deep ecosophy after all was more shallow or more muddy than you would expect from someone who warned that when you don't see the bottom of a lake, it might be because the water is muddy.' (Dag og Tid 004)

Bibliography

- J. T. Faarlund, S. Lie, and K. I. Vannebo. *Norsk referansegrammatikk*. Universitetsforlaget, 1997.
- C. Falk. *Non-referential subjects in the history of Swedish*. PhD thesis, Lund University, 1993.
- D. T. T. Haug. 'med' - the syntax and semantics of concomitance in Norwegian. In M. Butt and T. H. King, editors, *Proceedings of the LFG09 Conference*, pages 338–356. CSLI Publications, 2009.
- D. T. T. Haug. PROIEL Guidelines for Annotation, 2010. URL http://folk.uio.no/daghaug/syntactic_guidelines.pdf.
- J. B. Johannessen, L. Øverlid, K. Hagen, K. Kinn, and P. E. Solberg. Utgreiing om tekstkorpus til språkteknologisk forskning og utvikling. Technical report, Tekstlaboratoriet, ILN, UiO, 2011.
- H. Lødrup. The diversity and unity of the accusative with infinitive construction: a norwegian perspective. *Linguistics*, 46-5:891–916, 2008.
- E. Lundebj. *Om “utbrytningen”s opphav og innhold*, pages 280–301. Studentlitteratur, AB, 1976.
- A. Lynum, K. Hagen, J. B. Johannessen, and A. Nøklestad. Obt+stat: Evaluation of a combined cg and statistical tagger. In *NEALT Proceedings Series*, volume 14, 2011.
- E. Maier. Mixed quotation: between use and mention. In *Proceedings of LENS*, 2007.
- J. Nivre. Dependency grammar and dependency parsing. Technical report, Vxj University: School of Mathematics and System Engineering, 2005.
- J. Nivre. Incremental non-projective dependency parsing. In *Proceedings of NAACL HLT*, 2007.
- J. Nivre, J. Nilsson, and J. Hall. Talbanken05: A swedish treebank with phrase structure and dependency annotation. In *Proceedings of LREC*, pages 1392–1395, 2006.
- L. Tesnière. *Éléments de syntaxe structurale*. Klincksieck, 1965.