The LANCHART word order study

Working paper for the IC meeting 5th – 11th March 2008

Jensen, Torben Juel

Publication date:
2008

Document version
Publisher's PDF, also known as Version of record

The LANCHART word order study

Working paper for the IC meeting 5th – 11th March 2008

Torben Juel Jensen

Table of contents

Introduktion

Modern Danish distinguishes between to different word orders: The so-called *main clause word order* is characterized by a topological first position called the *fundament field*, which serves both as a topicalization device and, via the contrast filled vs. empty, as the expression of realis vs. non-realis meaning. Furthermore, main clause word order is characterized by the fact that sentence adverbials and negations are placed *after* the finite verb:

kommer	han	ikke	i morgen? (empty fundament field: non-realis)
comes	he	not	tomorrow

"doesn't he come tomorrow?"

han	kommer	ikke	i morgen (filled fundament field: realis)
he	comes	not	tomorrow

"he doesn't come tomorrow

i morgen	kommer	han	ikke (filled fundament field: realis)
tomorrow	comes	he	not

"tomorrow, he doesn't come"

In contrast, the so-called *subordinate clause word order* has no fundament field (the subject of the clause is always placed in the first position), and sentence adverbials are always placed *between* the subject and the finite verb:

han	sagde	*at*	*han*	*ikke*	*kommer*	*i morgen*
he	said	that	he	not	comes	tomorrow

"he said that he won't come tomorrow"

Especially in spoken Danish there is a great deal of variation in the word order of subordinate clauses, and subordinate clauses with main clause word order are rather frequent:

han	sagde	*at*	*han*	*kommer*	*ikke*	*i morgen*
he	said	that	he	comes	not	tomorrow

han	sagde	*at*	*i morgen*	*kommer*	*han*	*ikke*
he	said	that	tomorrow	comes	he	not

This variation is not new, it has existed for several centuries. Historically, the subordinate clause word order as expression of dependency is the innovation. A study by Frans Gregersen and Inge Lise Pedersen has demonstrated that the subordinate clause word order during the period 1500-1900 developed into being close to obligatory in subordinate clauses in *written* Danish. In spoken Danish, however, main clause word order is still rather frequent in the 20th century. Gregersen and Pedersen find that in (a sample of) the original BySoc-material, 35 % of the subordinate clauses have main clause word order. These subordinate clauses are predominantly nominal clauses introduced by *at* (English *that*) and causal clauses introduced by *fordi* (English *because*). A survey of (early) 20th century rural dialects indicates that main clause word order in subordinate clauses is much more common in Jutland dialects than in Zealand dialects, and that the Jutland dialects also have main clause word order in the largest number of clause types (Gregersen & Pedersen 2000).

It has been debated whether the variation documented in (at least) spoken modern Danish with regard to word order in subordinate sentences should be considered a sociolinguistic variable or as expressions of two different grammatical contents. Lars Heltoft and Tanya Christensen (proponents of the Danish Functional Linguistics school) have proposed that there is a semantic difference between "subordinate clause" and "main clause" word order which is in principle independent of the main/subordinate status of the clause. Main clause word order, which they label *declarative* word order, codes "assertive potential" – "informativity" - while subordinate clause word order (here labelled *neutral* word order) is neutral with respect to illocution. It is "unmarked" in that respect, and it can therefore be used both with and without assertiveness or informativity (Heltoft 1999 & 2005; Christensen 2006).

According to the generative grammarian Steen Vikner, subordinate clauses with main clause word order in modern Danish should simply be considered embedded main clauses (embedded V2). Main clause word order is thus only possible in "bridge verb contexts", e.g. as complements to the

verb *sige* (English *say*).[1] According to Vikner, true variation with respect to word order in subordinate clauses can only be found in medieval Danish, which had the possibility of V° to I°-movement in contrast to Danish after the 16th century. In modern Danish (and all other Scandinavian languages except Icelandic), the finite verb is always placed in V° (Vikner 1999 & 2004).

The LANCHART study of word order in subordinate clauses began last year and has, as of yet, only included the so called explorative sub-corpus and the informants from Copenhagen (BySoc 1). It is not definitively decided how much of the LANCHART corpus this study should include, but we expect in addition to BySoc to include at least one of the Jutland projects, i.e. Vinderup or Odder.

Digital marking and coding process

In order to study the variation in word order in subordinate clauses, *all* subordinate clauses containing a sentence adverb or a negation should be marked and coded for word order. This could of course be done by reading "manually" through the complete transcripts, but in order to speed up the process, AND to make it as consistent as possible, a semi-automatic approach has been chosen. Before the analytical coding begins, all text strings which may function as sentence adverbs are therefore marked automatically with a Praat script.[2] Afterwards, the human analytical coders[3] check all the marked sentence adverbial candidates, deciding in each case whether it actually is a sentence adverbial (many of the candidates are polysemous and can also function as conjunctions and/or content adverbials) and whether it occurs in a subordinate clause. In the cases where the marked adverbial candidate *does* function as a sentence adverbial in a subordinate clause (and is not placed

---

[1] There is some dispute as to which verbs allow embedded V2. EV2 is reported to be *ruled out* after factive verbs, after inherently negative verbs, after verbs introducing irrealis complements and after verbs in sentences that are negated, modalised or interrogative (Heycock et al. 2003).

[2] The script marks 107 different text strings which are all considered to be sentence adverbial candidates. The list of candidates has two sources: The starting point was a list of sentence adverbs compiled by Sanni Nimb as a result of her PhD-project on the syntax of Danish adverbs (Nimb 2004). This list contains 85 adverbs which Nimb in a study of *written* Danish has found only to occur in the fundament field or in the nexus field (i.e. before negations) according to Diderichsen's sentence schema. The other source was a small scale study of subordinate clauses in the "exploratory" sub-corpus (20 conversations from the LANCHART corpus selected so that the non-linguistic dimensions of variation in the LANCHART material would be present). In this study, we read through the conversations from beginning to end while coding *all* subordinate clauses. Afterwards, all subordinate clauses categorized as having "subordinate clause word order" or "main clause word order" (cf. page 10) - i.e. all clauses which are not neutral with respect to the opposition "main"/"subordinate clause" word order - and which did not contain any adverbs from Nimb's list were excerpted. All adverbials occurring in the nexus fields in these clauses were then evaluated and, if analyzed to function as sentence adverbials in at least one of their sense variants, added to the list of sentence adverb candidates.

in the "fundament field"), the beginning and end of the subordinate clause is marked by boundaries in a separate Praat tier aligned with the orthographic transcription:
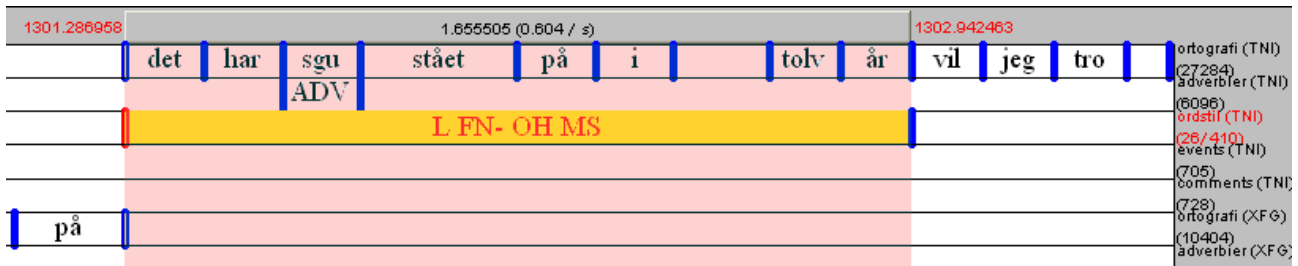

In the cases where one or more subordinate clause containing sentence adverbials occur embedded in another subordinate clause also containing sentence adverbials, all beginnings and ends are marked by boundaries, and embedment is marked by adding parentheses in the interval between the boundaries (in the "ordstil"-tier):
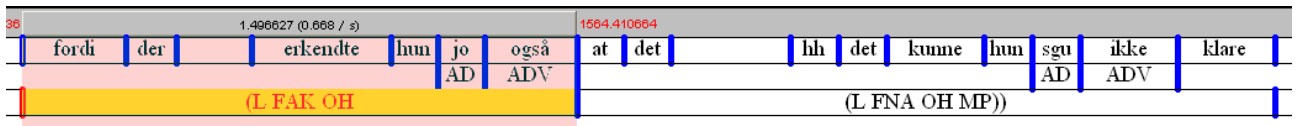
After marking the extent of the subordinate clause in the analytic Praat tier (the "ordstil"-tier), the clause is categorized with respect to *function*, *word order* and, as regards nominal clauses, *matrix clause*. Elliptic clauses and clauses which are broken off before an analyzable meaning has been expressed are given special codes (FE and FI, respectively) and not analyzed further. The codes are registered in the intervals aligned with the transcription of the clause (see the illustrations above). The diagram below illustrates the possible analytical codes, which are explained in the next section:

---

[3] The actual categorization and coding of the subordinate clauses occurring in the corpus is carried out by the student-assistants Stine Hartmann Bierre, Liva Hyttel-Sørensen, Rene Staustrup and Randi Sørensen under supervision of Torben Juel Jensen.

Coding scheme	Function (all subordinate clauses and fragments are coded)	Word order (only completed, non-elliptical clauses are coded)	Matrix clause (only nominal clauses are coded)
L	e.g. **F**NA	e.g. **O**B	e.g. **M**P
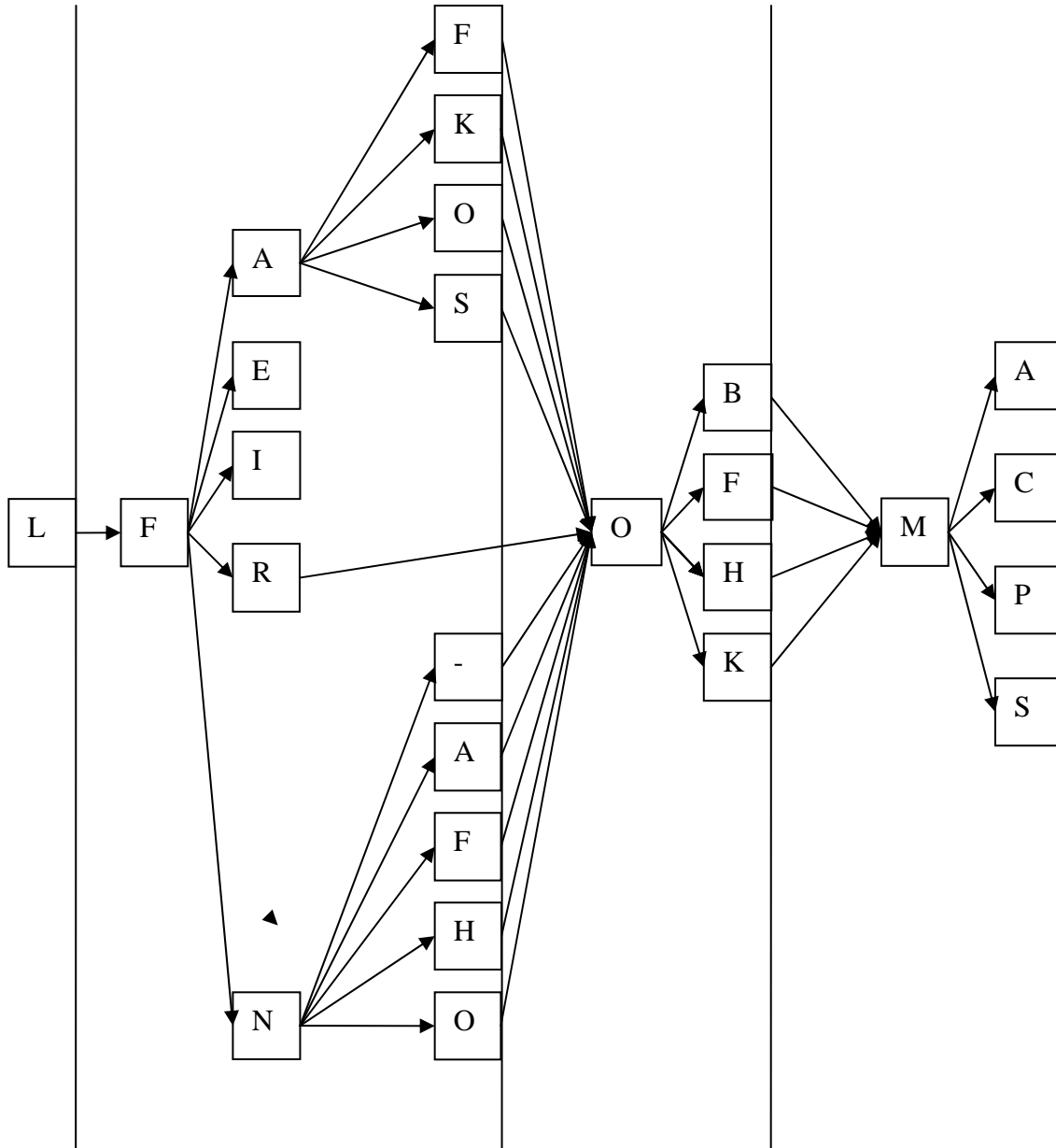
Analytical categories[4]

Function (F)

Adverbial subordinations (A)

Adverbial subordinations are subordinate clauses that have adverbial functions in the main clause. They specify circumstances regarding the content of the main clause such as time, place, cause, consequence, condition, contrast, manner, degree etc. The clauses are categorized into four variants according to their form:

FAF **"**Common conjunction" - there is no explicit subordinating conjunction in the subordinate clause as it is implied by the preceding (coordinated) subordinate clause, e.g.

- det er fordi   at   øh han vist   nok   øh har mødt en pige
  it   is because that uh he   probably enough uh has met   a   girl

- og øh familien   i hvert fald ikke mener at   han skal   giftes med med hende
  and uh the-family in any case not think that he   shall marry with with her

"it's because he has met a girl I think and uh the family does definitely not think he should marry her"

FAK Causal clause initiated with the conjunction **'**fordi**'** / **'**for det**'** (English 'because'), e.g.

- han var   træt fordi   den nat   havde han ikke sovet
  he   was tired because that night had   he   not   slept

"he was tires because he hadn't slept that night"

- bare for det   at   man er rig   behøver man jo   ikke have
  just for that that you are rich need   you after-all not have

- forstand   på penge
  reason   on money

"just because you're rich you don't necessarily have much money sense"

FAS "Interrogative" conditional clause, i.e. a conditional clause with no subordinating conjunction and with empty fundament field (V1 word order), e.g.

- bliver   det ikke regnvejr   i morgen tager vi   til   stranden
  become it   not rain-weather tomorrow go   we to   the-beach

"if it doesn't rain tomorrow we'll go to the beach"

---

[4] Thanks to Liva Hyttel-Sørensen who for helped me translate this section from the Danish version of the coding manual.

FAO   Other adverbial subordinations than the types described above, e.g.

    - og der boede jeg <u>indtil jeg nok</u>   var hh fire år
     and there lived I until I probably was hh four years

    "and I lived there till I was probably four years old"

    - <u>hvis det ikke bliver regnvejr   i morgen</u> tager vi på stranden
     if it not become rain-weather tomorrow go we to the-beach

    "if it doesn't rain tomorrow we'll go to the beach"

    - vandet var så koldt <u>at vi utvivlsomt frøs som små hunde</u>
     the-water was so cold that we undoubtedly froze like little dogs

    "the water was so cold that we without a doubt froze like little dogs"

Nominal subordinations (N)

Nominal subordinations (noun clauses) are subordinate clauses with substantival functions in the main clause. They typically function as (direct) objects, but they can also function as subjects, predicates and complements in prepositional phrases. Nominal subordinations can also be recognised by the fact that they are <u>syntactically obligatory</u>. The nominal subordinations are categorized into five variants according to their subordination conjunction:

FN-   No subordinating conjunction, e.g.

    - jeg tror <u>det der det der med at tage til Argentina det er snarere</u>
     I think that there that there with to go to Argentina it is rather

    - <u>en eller anden drøm</u>
     one or other dream

    "I think that whole thing about going to Argentina is probably just a dream"

FNA   Initiated by the subordinating conjunction 'at' (English 'that'), e.g.

    - familien mener <u>at han partout   skal giftes med med den pige der</u>
     the-family thinks that he at-any-cost shall marry with with that girl there

    "the family thinks he has to marry that girl"

FNF   **"**Common conjunction" - there is no explicit subordinating conjunction in the subordinate clause as it is implied by the preceding (coordinated) subordinate clause

    - jeg er glad for at det er godt vejr   <u>og det ikke regner</u>
     I am glad for that it is good weather and it not rains

    "I'm glad that the weather is fine and it doesn't rain"

FNH   Initiated by the subordinating conjunction 'om' (English 'if') or a 'hv'-word (equalling 'wh'-words in English), e.g.

- gad   vide   <u>om jeg ikke har   familie i   Argentina</u> ha
  liked know if   I   not have family in Argentina ha

"I wonder if I haven't got relatives in Argentina ha"

- det har jo   også været begrænset <u>hvad min mor   egentlig har</u>
  it   has as-you-know also been limited   what my mother really   has

  <u>vidst   om   det</u> ikke
  known about it not

"it has been limited how much my mother really has known about this, you know"

FNO   Initiated by other subordinating conjunctions than the ones described above.

Relative subordinations (including the subordinate clause in cleft-constructions) (R)

Relative subordinations are subordinate clauses that have one or more segments (sentence members) <u>in common</u> with the main clause. Their function is to present specifications to the common segment(s), either by reducing the number of possible referents (restrictive relative clauses) or by adding further information about the referent (parenthetical relative clauses). Relative subordinations are initiated by 'der' (English 'who'/'which'/'that'), 'som' (English 'who'/'which'/'that') or 'hv'-words, or occur without conjunction.

FR   - vi er   dem <u>de andre ikke må   lege med</u>
     we are   those the others not may play with

"we are the ones the others aren't allowed to play with"

- husker   du   den   sommer <u>hvor vi aldrig var   hjemme</u>
  remember you that summer where we never were home

"do you remember the summer where we were never at home?"

Word Order (O)

The word order in subordinations is determined only by the placement of <u>sentence adverbials and negations</u> (only clauses with sentence adverbials or negations are coded).

Subordinate clause word order

OB   A sentence adverbial is placed <u>between</u> the subject and the finite verb (s a v), e.g.

- <u>det er fordi   at   øh han vist   nok   øh har mødt en   pige</u>
  it   is because that uh   he probably enough uh has met   a   girl

"it's probably because he's met a girl"

Main clause word order

OH   A sentence adverbial is placed <u>after</u> subject and finite verb (v s a / s v a), e.g.

- <u>fordi   han havde jo   prøvet lidt   af hvert</u>
  because he   had   as-you-know tried   little of   every

"because he'd tried all sorts of things"

Fronted adverbial

OF   A sentence adverbial is placed <u>before</u> the subject. These clauses have the word order <u>a s v</u>. If the clause is initiated by a subordinating conjunction the adverbial is placed between the conjunction and the subject, e.g.

- <u>hvis hvis ikke man var   på cykel</u> så   gik   man bare før
  if   if   not   you were on bicycle then went you   just before

- hjemmefra
  from-home

"if you weren't going by bike then you just had to leave home earlier"

Sentence intertwining

Sentence intertwining is a special case of extraction where a non-WH constituent of a subordinate clause occurs in the first position of the matrix clause. As in cleft constructions, an 'empty slot' is left behind in the subordinate clause.

OK   - <u>det   tror   jeg ikke har   haft nogen indflydelse</u>
           that think I   not   have had   any   influence

"I don't think it has any influence"

Matrix clause (M)

Only nominal subordinations are coded according to their matrix clause.

MC   The matrix clause embeds <u>linguistically marked</u> *quoted/direct speech*. This may be indicated by the speaker distorting his voice (even when quoting himself) or by retaining the tense and person from the cited situation, e.g.

- så   sagde hun nu   kommer han sikkert   igen   det   kvaj
  so said   she now comes   he probably again that fool

"then she said now he's probably coming back, that fool"

- så   siger værkføreren bare at   der   er sgu   altid   så meget i
  then says the-foreman just   that there is [mild swear word] always so much in

- vejen   med Dem
  the-way with you [polite form]

"then the foreman just says that there's always something wrong with you"

MP   The matrix clause is *presupposing* (factive), marking the subordination as presupposed, i.e. presumed to be true. Presupposed subordinations don't change truth value if the matrix clause is negated – the information (the proposition) of the subordination is still presumed to be true, e.g.

- det er mærkeligt at   man ikke gør det
  it   is   strange   that you   not   do   it

"it's strange that you don't do it"

MS   The matrix clause expresses *secondary* information, i.e. it clearly constitutes the background while the content of subordinate clause is in the foreground. Consequently, the matrix clause has a qualifying function in relation to the subordinate clause, which is *semantically* the main clause, e.g.

- <u>mester har vist   skrevet den under</u> tror   jeg
  master has probably written it   under think I

"the boss has probably signed it, I think"

- jeg mener <u>du   kan da   bare tage lidt   tidligere hjem</u>
  I   mean you can then just   take little earlier   home

"I mean surely you can just go home a little earlier"

MA   Other types of matrix clauses than the ones mentioned above.

References

Christensen, Tanya Karoli (2006): *Hyperparadigmer – en undersøgelse af paradigmatiske samspil i danske modussystemer*. PhD thesis. Institut for kultur og identitet, Roskilde Universitetscenter.

Diderichsen, Paul (1974 [1946]): *Elementær dansk grammatik*. København: Gyldendal.

Gregersen, Frans & Inge Lise Pedersen (2000): ”A la Recherche du Word Order Not Quite Perdu”. In: Herring, Susan C. et al. (ed.): *Textual Parameters in Older Languages*. Amsterdam: John Benjamins Publishing Company.

Heltoft, Lars (1999): ”Hierarki og rækkefølge – skandinavisk ledstilling i funktionel grammatisk belysning”. In: Jensen, Per Anker & Peter Skadhauge (ed.): *Sætningsskemaet i generativ grammatik*. Kolding: Institut for Erhvervssproglig informatik og kommunikation, Syddansk Universitet.

Heltoft, Lars (2005): ”Ledsætning og letled i dansk. OV-sætningens rester”. In: Heltoft, Lars; Jens Nørgaard-Sørensen & Lene Schøsler (ed.): Grammatikalisering og struktur. København: Museum Tusculanum.

Heycock, Caroline; Anthony Alderson, Jonathan Brennan,Golda Fischer, Victoria Gall, Nicole Gregoire, Teo Juvonen, Tom Kelly, Sophie Krauss, Jennifer Pope, Sofia Sanchez-Romero (2003): *Verb movement and the philosopher's stone - Some comparative data from Faroese, Danish, and Swedish*. University of Edinburgh.

Nimb, Sanni (2004): *Danske adverbier mellem leksikon og syntaks*. PhD thesis. Institut for Nordiske studier og sprogvidenskab, Købehavns Universitet.

Vikner, Sten (1999): "Vo-til-Io flytning og personfleksion i alle tempora". In: *Íslenskt mál, 19, 81-128*.

Vikner, Sten (2004): ”Det finitte verbums position i dansk før og nu”. Peter Widell & Mette Kunøe (ed.): *10. møde om udforskningen af dansk sprog*. Århus.