

Prominence relations and a-structure - evidence from Danish synthetic compounding*

Bjarne Ørsnes
Center for Sprogteknologi
bjarne@cst.ku.dk

August 14, 1996

Abstract

This paper discusses the relevance of various prominence relations for the lexical process of synthetic compounding. In particular I will pay attention to the notion of argument structure and to the issue of how much syntax goes into a-structure. I will argue that previous approaches to synthetic compounding in English fail to account for the Danish data and that the process of compounding is sensitive to a hierarchy of grammatical functions as embodied in several linguistic frameworks (e.g. HPSG, LFG). I will therefore conclude that syntactic prominence along the lines defined in LFG must be lexically available. In LFG this means that the LMT component must apply in the lexicon associating each argument position with its grammatical function.

1 Introduction

The aim of this paper is to discuss various conceptions of argument structure on the basis of Danish synthetic compounding. This analysis was originally developed within the framework of Head-Driven Phrase Structure Grammar but the present proposal will concentrate on how the results can be integrated in Lexical-Functional Grammar and LMT. The main focus, however, is on the amount of syntax encoded in argument structure. First I will briefly discuss a recent configurational approach to synthetic compounding ([Lieber, 1992]) and subsequently I will review different conceptions of a-structure (viz. [Rappaport and Levin, 1988, Grimshaw, 1990,

*Thanks are due to Nancy L. Underwood for removing the worst errors in my English.

Bresnan and Zaenen, 1990]) with a view to synthetic compounding. The order of presentation will be in terms of an increasing ‘syntactification’ of a-structure.

2 Prominence Relations

Current linguistic theory recognizes three distinct dimensions of prominence ([Bresnan, 1995]).

- **Thematic prominence:** Thematic prominence is a relation holding between arguments belonging to the same argument structure list according to a universal hierarchy of thematic roles which says that certain roles are more prominent or topical than other roles.
- **Syntactic Rank:** Syntactic rank is a relation between syntactic arguments of a predicator defined over a hierarchy of grammatical functions. According to this principle, subjects are more prominent than objects, and objects more prominent than obliques.
- **Configurational prominence:** The notion of configurational prominence is a relation between phrase structure positions and may be stated in terms of linear precedence or c-command relations.

3 The problem: Incorporation in synthetic compounds

By the term incorporation I refer to the phenomenon of a stem exhibiting dual behaviour: it is a syntactic argument of a predicator, but morphologically part of that predicator (rephrased after [Mohan, 1995]).¹ I will, however, concentrate exclusively on a subclass of incorporation structures, namely the one where the predicator is an argument taking deverbal nominal.² I will term this subclass Complex Event Compounds (CECs) after [Grimshaw, 1990]). A CEC is consequently a complex event-denoting compound with a deverbal argument taking head. Apart from being argument taking, these constructs have a couple of syntactic characteristics:

- They do not pluralize:
(1) **papirfremstillinger*
PAPER PRODUCTIONS
- They occur with an optional definite determiner but not with an indefinite determiner:

¹Note that this definition excludes the incorporation of adjuncts which are not syntactic arguments of the predicator. In this paper I will not discuss the incorporation of adjuncts even though I will point to some difficulties associated with adjunct incorporation and an LFG account of synthetic compounding in Danish.

²Furthermore I will only discuss compounds with the suffix *-ing*, which is the only productive nominalizing affix in modern Danish.

(2) *papirfremstillingen*
THE PAPER PRODUCTION

(3) **en papirfremstilling*
A PAPER PRODUCTION

- They occur with a case marking preposition:

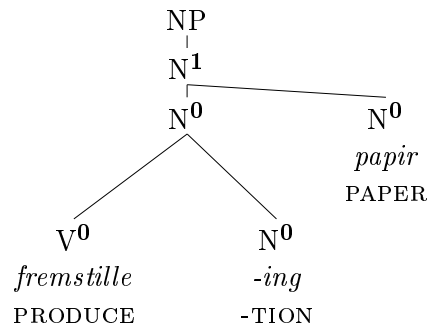
(4) *flisebelægningen på gaden* (result reading)
THE FLAG COVERING ON THE STREET

(5) *flisebelægningen af gaden* (event reading)
THE FLAG COVERING OF THE STREET

4 Configurational Prominence

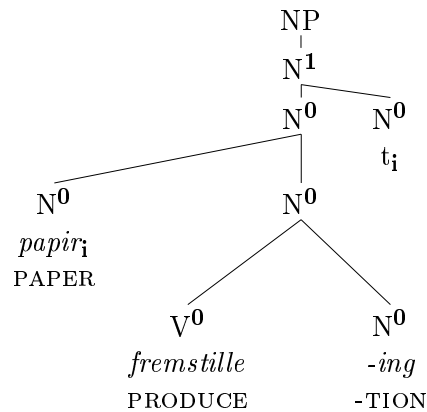
[Lieber, 1992] develops an account of deverbal compounds within the framework of GB.³ She assumes that all deverbal compounds have a syntactic deep-structure, where the incorporated element appears in its canonical phrase-structure position where it is assigned a theta-role by the predicator. This is possible since she assumes a modified version of X-bar theory which allows lexical non-heads as depicted in the figure below:

d-structure representation for *papirfremstilling* ('paper production'):



Since, however, a governor cannot assign case to a non-maximal projection this element has to move inside the word, where case assignment is not required yielding the following s-structure:

³As Lieber notes herself her account makes the same predictions as earlier accounts of English synthetic compounding despite the revised theoretical underpinnings ([Lieber, 1983, Roeper and Siegel, 1978]).



Here the direct argument of the verb is adjoined to the head-noun leaving a coindexed trace outside the word. This approach has some obvious advantages:

- The lack of phrasal non-heads follows from an independent principle: the case theory.
- The generic interpretation of the non-head follows from the fact that a lexical element lacks a determiner.
- We have a unified account of theta-role assignment which is always to the right under adjacency.

This approach, however, makes a couple of predictions which renders it inadequate in view of the Danish data, namely that only elements directly theta-marked by the predicator, in effect direct objects may incorporate. External arguments (subjects) and indirect arguments (obliques) are excluded.

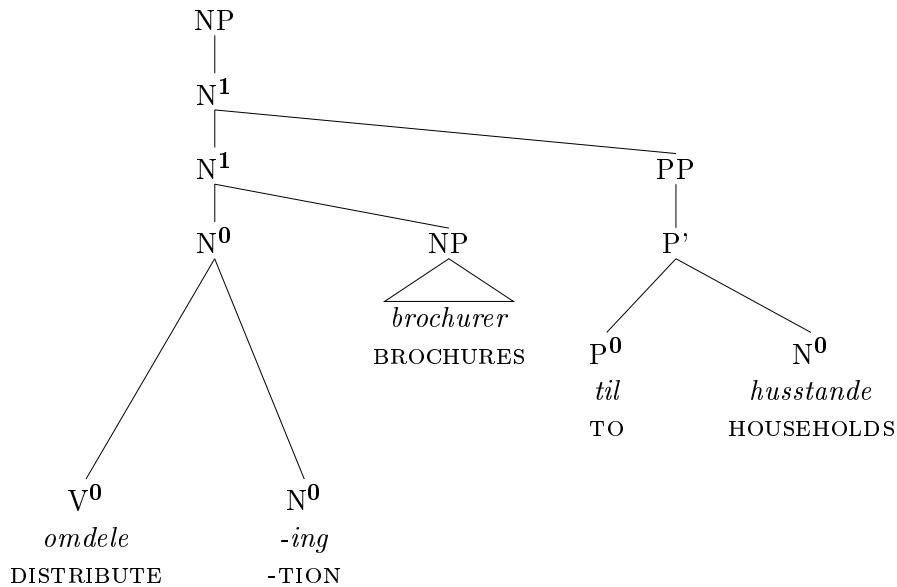
Let us first take a look at a nominalization with two internal arguments. The fairly common compound *husstandsomdeling* ('household distribution') as in:

- (6) *husstandsomdeling af brochurer*
HOUSEHOLD DISTRIBUTION OF BROCHURES

is based on a nominalization with two internal arguments:

- (7) *omdeling af brochurer til husstande*
DISTRIBUTION OF BROCHURES TO HOUSEHOLDS

Since a predicator can only case and theta-mark one argument I assume the following d-structure for this compound:

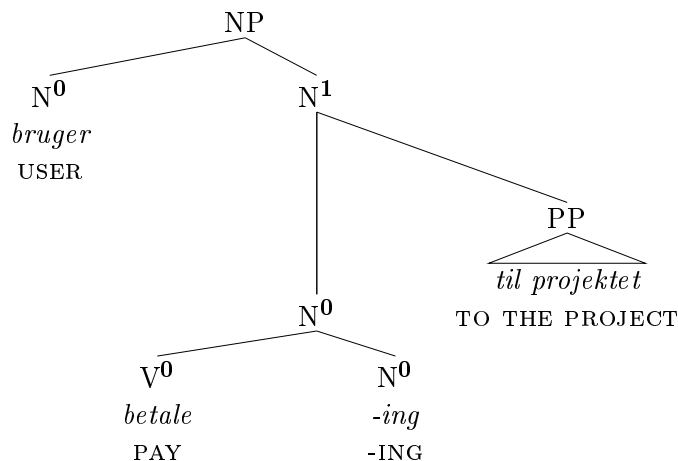


Again we may assume that the N^0 *husstande* ('households') is forced to move since the governing preposition cannot assign case to a non-maximal projection. But this leaves us with a preposition which never appears in the final derivation of the word. We would have to assume that this preposition is deleted at some level of representation but in that case the trace left behind by the moved element would lack a theta-role assigner and it would not be properly governed in violation of the Empty Category Principle.

Turning to the incorporation of external arguments, we find the well-established compound:

- (8) *brugerbetaling til projektet*
 USER PAYMENT TO THE PROJECT

with an incorporated external argument (subject). For this compound I assume the following d-structure representation:



Also in this case, the external argument *bruger* ('user') is forced to move since it is a lexical element to which no case can be assigned, but the trace left behind by the moved element will not be properly governed either, and furthermore the trace will asymmetrically c-command its antecedent, which is illegal. As a matter of fact this boils down to the prediction that intransitive, unergative verbs cannot occur in incorporated structures at all, which is counter-exemplified by the following CECs:

(9) *larvevandring*
LARVA MIGRATION

(10) *politipatroljering*
POLICE PATROLLING

Incorporation of the sole argument of unergatives is just as possible as the incorporation of unaccusative arguments:

(11) *ammoniakfordampning*
AMMONIA EVAPORATION

There is thus no way how a movement analysis along the lines suggested by [Lieber, 1992] can be reconciled with the Danish data, and configurational prominence does not seem to give an adequate account of Danish synthetic compounding.⁴

5 Notions of a-structure:

In view of the shortcomings of the purely configurational approach discussed above, I will instead turn to the lexical level of argument structure. In the following I will briefly review three conceptions of a-structure and discuss to what extent they manage to account for the Danish data.

5.1 Thematic Prominence

[Grimshaw, 1990] assumes that a-structure consists of a number of syntactic variable positions, ordered in accordance with a universal hierarchy of thematic roles as depicted below:

AGENT \ll EXP \ll GOAL/SOURCE/LOCATION \ll THEME

Grimshaw does not consider these thematic roles to be primitives but rather to be labels identifying specific arguments positions. Grimshaw further assumes that theta-marking respects this hierarchical ordering in the sense that roles lower on the hierarchy are discharged before roles higher on the hierarchy and further she assumes that theta-marking in morphological structures precedes theta-marking in the syntax. Thus, the specific prediction with respect to compounding is:

⁴Apart from the fact that more oblique elements in configurational languages tend to more deeply embedded than less oblique elements [Bresnan, 1995] and, as we will see later, more oblique elements tend to incorporate before less oblique ones.

... when the head takes more than one internal argument, the least prominent must be inside the compound and the more prominent outside.
(pp. 14-15)

This principle accounts for the contrast between:

(12) *gift-giving to children* (THEME incorporated)

(13) **child-giving of gifts* (GOAL incorporated)

Let us take the Danish verb *behandle* ('to treat') as an example. In its medical sense this verb has (at least) 3 arguments, which are ordered as depicted below:⁵

<i>behandle</i>	(x	(y	(z))
TREAT	agent	instr	theme

From the principle above we should predict that compounds headed by *-behandling* ('treatment') would occur with themes as their non-heads. And yet it turns out that by far the most compounds headed by *-behandling* ('treatment') contain instruments, and furthermore the less prominent argument is realized outside the compound, contrary to Grimshaw's prediction:

(14) *metadonbehandling af narkomaner*
METHADON TREATMENT OF DRUG ADDICTS

(15) *varmebehandling af mælk*
HEAT TREATMENT OF MILK

We even find incorporated agents which are otherwise taken to be the most prominent thematic roles:

(16) *lægebehandling af narkomaner*
DOCTOR'S TREATMENT OF DRUG ADDICTS

From this we may conclude that synthetic compounding in Danish does not respect thematic prominence and that thematic prominence from this point of view should not be the structuring device for a-structure.

5.2 Manner of theta-role assignment

In between Grimshaw's more semantic conception of argument structure and the more syntactic approach of the Lexical-Mapping Theory (LMT) of LFG, we have the approach developed by [Rappaport and Levin, 1988]. Levin and Rappaport consider argument structure to be an explicit representation of the hierarchical relations between a predicator and its arguments. These hierarchical relations, however, are not based on a thematic hierarchy but rather on the manner of theta-role assignment, that is how an argument is assigned its thematic role. Levin and Rappaport distinguish three ways that an argument is assigned a thematic role:

⁵Grimshaw does not have an instrument in her hierarchy, but I assume that it counts as more prominent than theme.

1. By predication \Rightarrow External argument
2. Directly by predicator \Rightarrow Direct internal argument
3. By a preposition \Rightarrow Indirect argument

For the Danish verb *forske* ('to do research'), which takes a prepositional complement (an indirect argument) we have the following a-structure:⁶

forske i noget X < P_{theme} Y >
 TO DO RESEARCH INTO STH.

For the verb *udforske* ('to explore sth.') we have a direct internal argument, indicated with underlining:

udforske noget X < Y >
 TO EXPLORE STH.

Levin and Rappaport claim that certain morphosyntactic processes such as adjectival passive formation are sensitive to the manner of theta-role assignment in the sense that only direct internal arguments may become the external argument of the corresponding adjectival passive. This accounts for the following contrast:

(17) **det forskede emne*
 THE RESEARCHED SUBJECT

(18) *det udforskede emne*
 THE EXPLORED SUBJECT

But it turns out that the internal arguments (i.e. direct internal and indirect arguments) behave alike wrt. compounding as can be seen from the examples below:

(19) *genforskningen*
 GENE RESEARCH

(20) *havudforskningen*
 SEA EXPLORATION

Compounding is obviously not sensitive to the specific manner of theta-role assignment. This does not mean in itself that Levin & Rappaport's conception of a-structure is not appropriate for the purpose of compounding. The manner of theta-role assignment ultimately determines the grammatical functions born by these variables in the syntax and these a-structures reflect an ordering which is familiar from other linguistic frameworks (e.g. HPSG). But still one could maintain that the explicit representation of the manner of theta-role assignment is irrelevant for the purpose of compounding, being superseded by an obliqueness ordering of grammatical functions.

⁶It is a little unclear how the preposition *i* 'in' comes to assign a theme-role but I will not discuss that matter here.

5.3 Intrinsic Syntactic Classification

Let us now turn to an even more syntactic approach to a-structure, namely the one of the Lexical-Mapping Theory (see e.g. [Bresnan and Kanerva, 1992, Bresnan and Zaenen, 1990]). In addition to a more or less familiar hierarchy of semantic roles LMT assumes a set of syntactic under-specification features regulating the mapping of semantic arguments to grammatical functions where the grammatical functions are decomposed into the features [+/-o(bjective)] and [+/-r(estricted)]. The semantic arguments in the a-structure are intrinsically classified for their ultimate syntactic function. Patient-like roles are classified as [-r] thus mapping to subject and object while other roles are classified as [-o], thus mapping to either subject or oblique.

The question is whether these intrinsic classifications tell us anything about compounding. It has been proposed by [Alsina, 1990] (quoted after [Markantonatou, 1994]) that roles more prominent than that of instrument are intrinsically classified as [-o] and that non-objective arguments cannot incorporate. Considering the sample verb *behandle* ('treat') we have the following a-structure:

<i>behandle</i>	<	ag	instr	th	>
TREAT					
intrinsic		[-o]	[-o]	[-r]	

Contrary to Alsina's principle instrument has been assigned the intrinsic classification [-o] since instruments usually map to oblique phrases. In any case the generalization that a [-o] does not incorporate cannot be maintained since we have a variety of compounds with incorporated non-objective arguments (cf. (14) and (15)). Furthermore such a principle says nothing about the behaviour of other arguments of a predicator in compounding.

We may then wonder if the restriction should be restated in terms of the [-r]-feature, so that [-r] elements may not incorporate. But this would exclude agents and themes, which otherwise abound in compounds. And finally, by the same token, if we state the generalization in terms of the [+o]-feature, we have in effect said nothing about a-structures containing no [+o]-arguments at all. This would ultimately imply either that no generalizations can be made for a-structures without a [+o]-argument or that we miss a generalization.

Secondly we may define an obliqueness ordering on the basis of the syntactic under-specification features and assume that compounding is sensitive to this kind of obliqueness ordering:⁷

(21) [-o] > [-r] > [+o]

If, however, we were to define the generalizations about compounding in terms of obliqueness of these features we would predict that [-r] elements incorporate prior to [-o] arguments. This prediction is not correct for Danish since the syntactic under-specification features are not fine-grained enough to distinguish subjects and obliques

⁷This possibility has been suggested to me by Joan Bresnan (p.c.).

of the same a-structure. Obliques consequently count as more prominent than objects, even though the hierarchy of grammatical functions has it the other way around. Thus in the example with *behandling* ('treatment'), discussed above, a [-o]-argument (i.e. the instrument) tends to incorporate and not the [-r]-argument even though the obliqueness ordering in (21) above would predict otherwise.

6 Between syntax and a-structure: Hierarchy of grammatical functions

It may seem from the foregoing discussion that compounding is totally unrestricted in Danish, but that does not seem to be the case. As a matter of fact compounding seems to be sensitive to the obliqueness ordering of grammatical functions.

The hierarchy of GFs takes the familiar format depicted here (see e.g. [Bresnan, 1995]):

SUBJ ≪ OBJ ≪ OBJ2 ≪ OBL ≪ COMPL

A subject is the least oblique element, and then we have an order of increasing obliqueness down to predicative and sentential complements (COMPL). For the nominalization *behandling* ('treatment') we consequently have the following subcategorization frame ordered in accordance with the functional hierarchy:

behandling <SUBJ OBJ OBL>
TREATMENT

In compounding we may then pick the least oblique element, the subject and make it part of a compound leaving the remaining internal arguments to be realized outside the compound:

(22) *lægebehandling af narkomaner med metadon* (SUBJ)
DOCTOR'S TREATMENT OF DRUG ADDICTS WITH METHADON

Alternatively we may choose the most oblique element and turn it into a non-head, while we cannot pick the element in between and still have a more oblique element outside the compound.

(23) *metadonbehandling af narkomaner* (OBL)
METHADON TREATMENT OF DRUG ADDICTS

(24) *??/* narkomanbehandling med metadon* (OBJ)
DRUG ADDICTS TREATMENT WITH METHADON

The last example is marginal at best.

The generalization about compounding in Danish may then be stated as follows:

1. In compounding, the most oblique element of the COMPS list may be incorporated.

If the most oblique element is optional, the second most oblique element may be incorporated. In that case the most oblique one can no longer be realized.

2. In compounding, the member of the SUBJ list may be incorporated.

If we look at the examples which were problematic for Levin and Rappaport's conception of a-structure, repeated in (26) and (28) below, we see that they receive a straightforward account, since we abstract away from the particular grammatical functions of the non-heads and simply note that their status as most oblique element makes them eligible for compounding.

(25) *forskning* < SUBJ OBL >
RESEARCH

(26) *genforskning*
GENE RESEARCH

(27) *udforskning* < SUBJ OBJ >
EXPLORATION

(28) *havudforskning*
SEA EXPLORATION

7 Integration of this analysis into LFG

As noted in the introduction LFG explicitly recognizes syntactic rank as a separate dimension of prominence. But this prominence is defined at the level of f-structure and is thus not lexically available for morphological processes such as compounding.⁸ Thus the question is whether syntactic prominence relations should be defined at f-structure or it should be present already at a-structure for the purpose of lexical processes such as the one considered here. One possibility would be for a-structure to reflect syntactic prominence. In that case, however, a-structure would no longer respect the hierarchy of thematic prominence.

The second solution would be for the LMT component to apply in the lexicon, in effect projecting nuclear f-structures within the lexical entries themselves. In that way syntactic prominence will be lexically available. This approach, however, fails to account for the incorporation of adjuncts which - although (presumably) the least oblique grammatical function - are not part of the nuclear f-structure. In order to provide for a lexical account of synthetic compounds with incorporated adjuncts, adjuncts must be introduced in the argument structure list of the lexical items, e.g. by means of a lexical rule as proposed by [van Noord and Bouma, 1994] for an HPSG treatment of adverbial scope and word order.

8 Conclusion

An examination of incorporation in Danish synthetic compounds shows that syntactic prominence is not that easily attributed to the level of syntactic f-structure and

⁸Contrary to HPSG where the obliqueness ordering is respected in the ARG_S feature and in the valency lists SUBJ and COMPS (see e.g. [Manning and Sag, 1995]).

that thematic prominence is not simply attributed to the lexical level of a-structure. Certain morphological processes applying in the lexicon require access to syntactic information and consequently syntactic information must be lexically available. Even though LFG does allow for a fairly straightforward integration of this analysis further work remains to be done regarding the treatment of adjuncts and the application of the LMT component in the lexicon.

References

- [Alsina, 1990] Alsina, A. (1990). Where’s the Mirror Principle. Ms. Stanford University - Unfortunately this paper has not been available to me, I have quoted after Markantonatou 1994.
- [Bresnan, 1995] Bresnan, J. (1995). Linear Order, Syntactic Categories and Empty Categories: On Weak-Crossover. In Dalrymple, M., Kaplan, R. M., Maxwell, J. T., and Zaenen, A., editors, *Formal Issues in Lexical-Functional Grammar*, pages 241–274. CSLI.
- [Bresnan and Kanerva, 1992] Bresnan, J. and Kanerva, J. M. (1992). Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. In Stowell, T. and Wehrli, E., editors, *Syntax & Semantics 26: Syntax and the Lexicon*, pages 53–101. New York: Academic Press.
- [Bresnan and Zaenen, 1990] Bresnan, J. and Zaenen, A. (1990). Deep Unaccusativity in LFG. In Dziwirek, K., Farrell, P., and Meijas-Bikandi, E., editors, *Grammatical Relations: A Cross-Theoretical Perspective*, pages 45–57. Stanford: CSLI/Stanford University.
- [Grimshaw, 1990] Grimshaw, J. (1990). *Argument Structure*, volume 18 of *Linguistic Inquiry Monographs*. MIT.
- [Lieber, 1983] Lieber, R. (1983). Argument Linking and Compounds in English. *Linguistic Inquiry*, 14(2):251–285.
- [Lieber, 1992] Lieber, R. (1992). *Deconstructing Morphology*. Chicago: Chicago University Press.
- [Manning and Sag, 1995] Manning, C. and Sag, I. A. (1995). Dissociations between Argument Structure and Grammatical Relations. ms.
- [Markantonatou, 1994] Markantonatou, S. (1994). Modern Greek Deverbal Nominals: An LMT Approach. ms., Essex: University of Essex.
- [Mohanan, 1995] Mohanan, T. (1995). Wordhood and Lexicality: Noun Incorporation in Hindi. *Natural Language and Linguistic Theory*, 13:75–134.
- [Rappaport and Levin, 1988] Rappaport, M. and Levin, B. (1988). What to Do with Theta-Roles. In Wilkins, W., editor, *Syntax & Semantics*, volume 21, pages 7–36. Academic Press, California.
- [Roeper and Siegel, 1978] Roeper, T. and Siegel, M. E. (1978). A Lexical Transformation for Verbal Compounds. *Linguistic Inquiry*, 9(2):199–260.

[van Noord and Bouma, 1994] van Noord, G. and Bouma, G. (1994). Adjuncts and the Processing of Lexical Rules. In *COLING-94*, volume 1, pages Kyoto, Japan: 250-57.