

An HPSG Account of Danish Pre-nominals

Anne Neville
Center for Language Technology
Copenhagen University
anne@cst.ku.dk

Abstract

This article addresses the issue of selection restrictions for noun phrase specifiers. Danish data is presented which shows that definiteness plays an important role in this respect. It is pointed out that an analysis is required in which the specifier, when present, leaves a mark on the projected phrase. This is achieved by assuming that specifiers are syntactic heads of noun phrase constructions. Further an elaborate classification of specifiers is also needed in terms of which selection restrictions may be formulated, along with a cross-categorial definiteness feature. These properties are part of the analysis proposed in this analysis.

Introduction

When investigating empirical data it becomes clear that noun phrases often have multiple specifiers appearing before the noun. An important goal of noun phrase analysis is the specification of selection restrictions for noun phrase specifiers and pre-nominals in general to account for combinations of specifiers. It is this goal that is pursued in this article.

In section 1 a set of Danish noun phrases are presented which form the basis of a discussion of what properties determine the restrictions on combinations of pre-nominals. In section 2 a number of previous HPSG analyses of noun phrases and pre-nominals are discussed. In section 3 the proposed analysis is introduced and sample analyses are shown. The proposed analysis has been implemented in the LKB system (Copestake 1999). A test suite consisting of the data in section 1 has been run and the results are presented in section 4. The article is concluded in section 5.

1 Some data

To narrow the focus of this article two sets of noun phrases have been selected which illustrate the importance of the feature of definiteness in the account of selection restrictions. One set consists of noun phrases with multiple definite specifiers,

as shown in (1), and the other set contains pre-nominals showing the dependence between specifiers and adjectives wrt. definiteness, as shown in (2).

- (1) a. *denne min begejstring*
(this my enthusiasm)
- b. *denne den sidste Rabbit-bog*
(this the last Rabbit book)
- c. *jeres den gamle grammofon*
(your the old record player)
- d. *deres den nærmeste nabo*
(their the nearest neighbour)
- e. *hende den tossede malerinde*
(her the crazy painter)
- f. *ham den forsvundne dreng*
(him the missing boy)
- (2) a. *de tre mindste skoler*
(the three smallest schools)
- b. **tre mindste skoler*
(three smallest schools)
- c. *de mange smukke ting*
(the many beautiful things)
- d. **de mange smukkeste ting*
(the many most beautiful things)

All the examples in (1) contain two definite specifiers. In (1a) the demonstrative specifier *denne* precedes a possessive specifier. In (1b) the demonstrative specifier *denne* precedes the definite article. In (1c) and (1d) the possessive specifiers *jeres* and *deres* likewise precede the definite article. Finally, in (1e) and (1f) the personal pronoun specifiers *hende* and *ham* precede the definite article¹. To account for these combinations, ruling out all other combinations, a detailed classification of definite specifiers is required. In (2) definiteness determines the possibility of combining specifiers and adjectives. In (2a) the specifier *tre* (three) combines with the definite adjective *mindste* (smallest). However, the unacceptability of (2b) indicates that this combination is licensed by the presence of the definite article. (2c) and (2d) show that *mange* (many) does not allow a following definite adjective, irrespective of the presence of a preceding definite article. Thus definiteness is not confined to the description of specifiers, but it is also relevant to the analysis of adjectives.

2 Previous analyses

A number of analyses of pre-nominals have been proposed within the framework of HPSG. Pollard & Sag (1994) propose a noun phrase analysis which is an NP

analysis for English. Pre-nominals are divided into specifiers and modifiers, and one specifier is allowed in an NP. The specification of selection restrictions consequently does not become relevant². A series of alternative analyses have been put forward since (Netter 1994, Kolliakou 1995, Allegranza 1998, Kathol 1998 and Börjars 1994).

Netter (1994) proposes a "DP" analysis for German. DP is in quotes because the analysis is not a DP analysis in the traditional sense (Abney 1987 and Delsing 1993). It is a DP analysis in the sense that the determiner is the head. However, determiners and nouns are assumed to be subtypes of a common nominal supertype. This means that both a noun phrase with a determiner and a noun phrase without a determiner may function as a maximal nominal phrase. Netter's analysis allows for multiple specifiers, in theory. However, he makes no attempt to specify selection restrictions.

Kolliakou (1995) also proposes a "DP" analysis, for Greek. It is a DP analysis in the same sense as Netter's analysis. However, Kolliakou's nominal type hierarchy is much more detailed than Netter's, and in addition to determiner and noun types she proposes demonstrative, numeral and adjective, all as subtypes of a common nominal supertype. Her analysis allows for multiple specifiers, and she specifies selection restrictions for them in terms of the nominal type hierarchy. Significantly, her analysis covers quantifying specifiers. Kolliakou's analysis covers a wide range of noun phrases. Her hierarchy is developed for Greek, and does not account for the Danish data, though.

Allegranza (1998) puts forward an "NP" analysis for Italian. Here NP is in quotes because it is not an NP analysis in the traditional sense (Chomsky 1970 and Jackendoff 1977). The noun is the syntactic head, but he introduces a marking feature by way of which the specifier non-head leaves a mark on the projected noun phrase. The value of the marking feature is a nominal type hierarchy as in the above-mentioned analyses. In his analysis determiner is a common supertype of a number of nominal subtypes. Allegranza's analysis likewise accounts for multiple specifiers, and selection restrictions are based on the determiner type hierarchy. His analysis also covers quantifying specifiers. Like Kolliakou's analysis, Allegranza's analysis covers a wide range of noun phrases. His analysis is developed for Italian, and again does not account for the Danish data without modifications.

Kathol (1998) proposes another "DP" analysis for English. Kathol also bases his analysis on a nominal type hierarchy, where determiner and nouns are subtypes of a common type. But like Pollard and Sag's analysis, his analysis only allows for one specifier, which makes it unable to account for multiple specifiers.

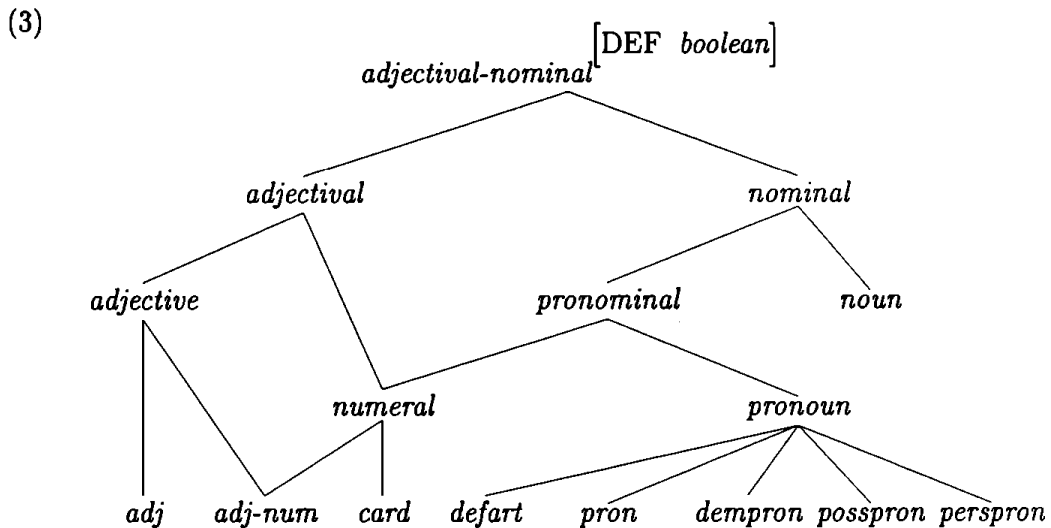
Finally, Börjars (1994) proposes an analysis very similar to Pollard and Sag's account. Consequently it has the same drawbacks as their account. However, it is interesting because it introduces definiteness as a syntactic primitive. It has already been shown how definiteness plays an important role in specifying selection restrictions. Börjars does, however, not explore the full potential of the feature.

The analyses referred to here serve to show that two properties of noun phrase analysis are important. Firstly, an analysis is required in which the specifier, when present, leaves a mark on the projected phrase. This can be achieved either by assuming that the specifier is the syntactic head, or by introducing a marking feature by way of which the specifier marks the projected noun phrase. Secondly, an elabo-

rate classification of specifiers is needed in terms of which selection restrictions may be formulated. Börjars' account further supports the observation that definiteness is a key feature in the analysis of noun phrases.

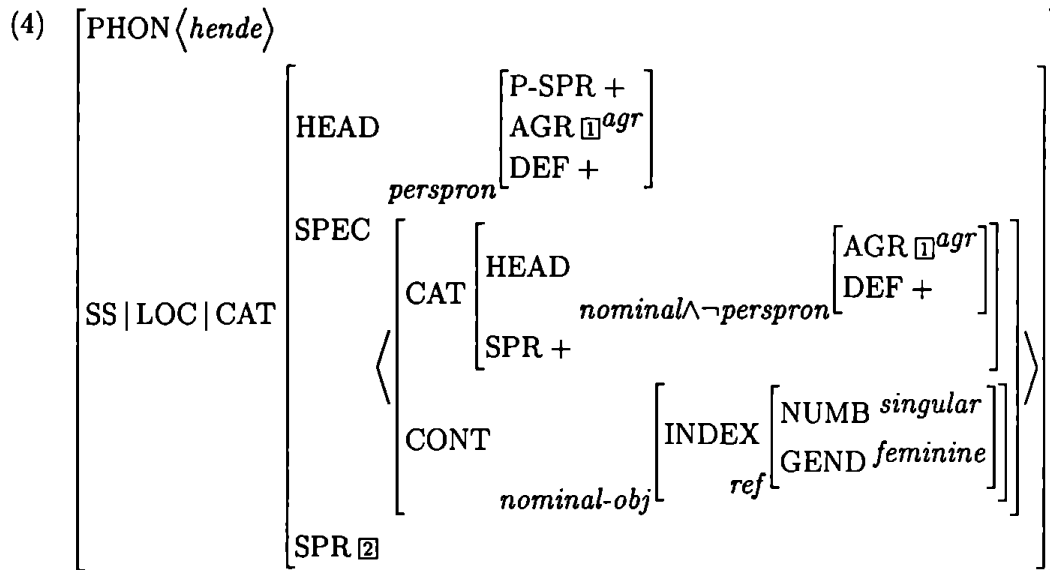
3 Proposed analysis

The analysis here builds on ideas from the accounts presented in section 2. A division into modifier and specifier pre-nominals is assumed where specifiers are analyzed as heads selecting their non-head sister. It is based on a subtyping of the HPSG *head* type into a hierarchy of adjectival and nominal types. It sets itself apart in a number of ways. First and foremost the analysis is distinguished by its emphasis on definiteness. This is reflected by the detail with which definite pronouns are subtyped, and by the adoption of a separate feature of definiteness pertaining to all adjectival and nominal categories. The subtyping of *adjectival-nominal* is shown in (3).



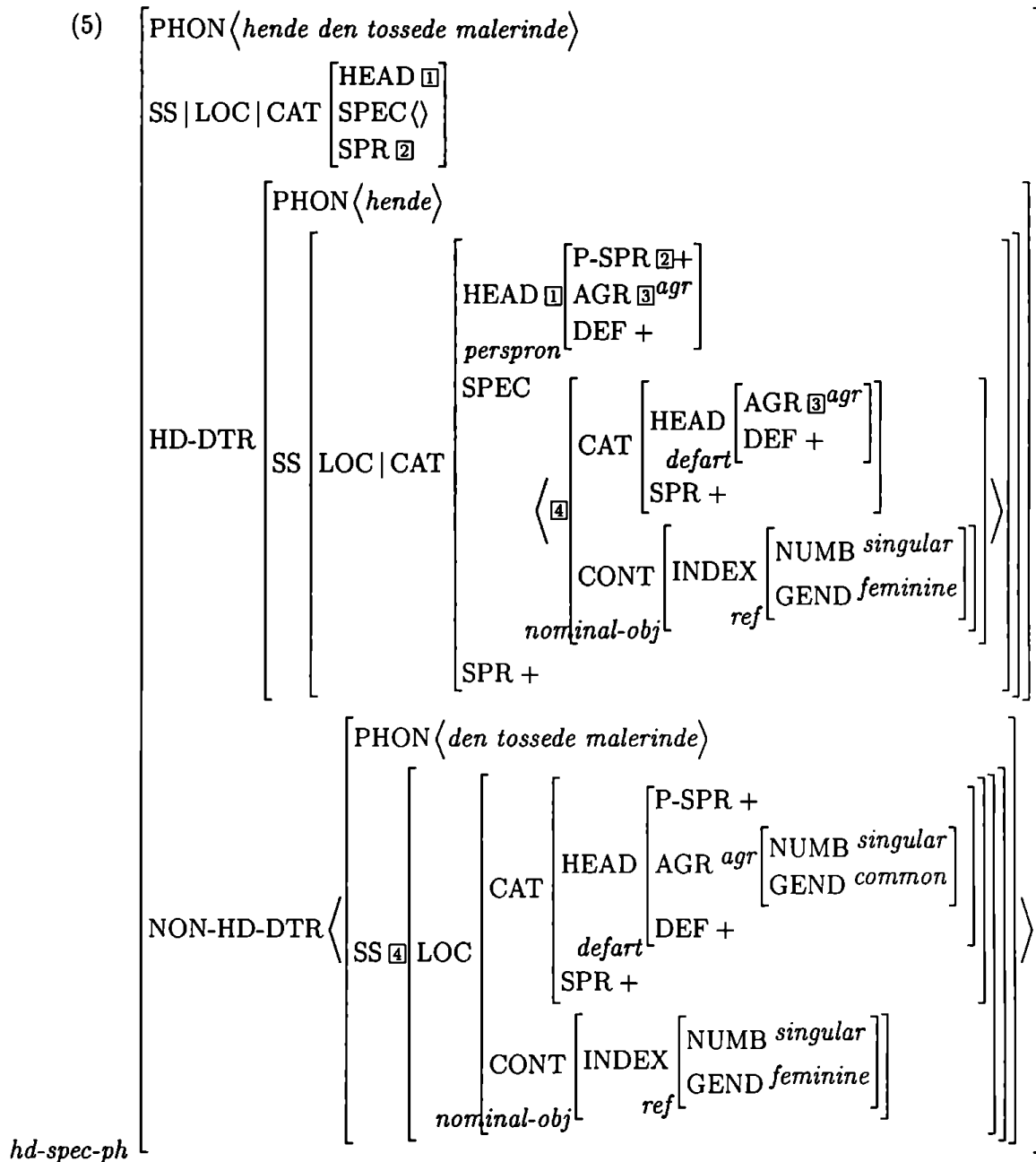
The attribute DEF is defined for the *adjectival-nominal* type, which means that it is inherited by all the subtypes of *adjectival-nominal*. The type hierarchy and the assumption that specifiers are syntactic heads of noun phrase constructions enable an account of the data in section 1. In the following two sample analyses are presented. The sample noun phrases are *hende den tossede malerinde* (her the crazy painter) and *tre mindste skoler* (three smallest schools).

In (4) the lexical entry and thus selection restriction for the personal pronoun specifier *hende* (her) is shown.



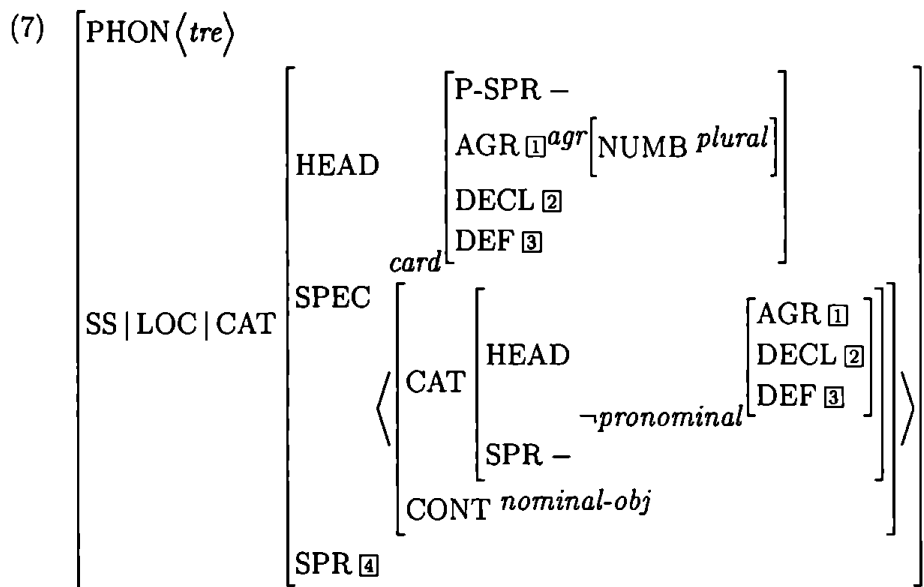
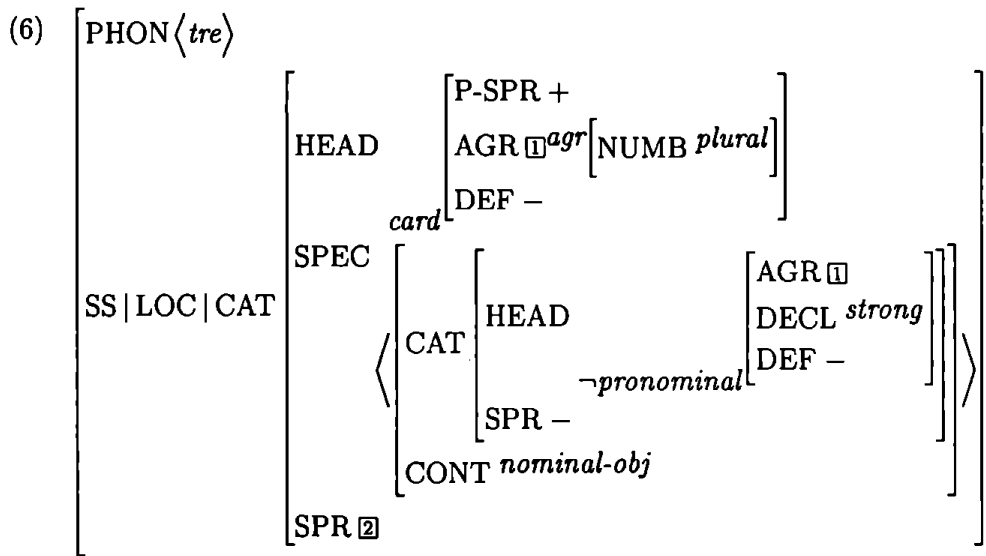
The value of HEAD in (4) shows that *hende* is of type *perspron* and DEF+. The P-SPR attribute indicates whether *hende* projects a maximal noun phrase or not. (4) further shows that *hende* is a specifier, the value of the SPEC attribute is a list restricting its non-head selectee. The selection restriction is that the selectee must be a *nominal*, but not a *perspron*. In addition it must be DEF+ and SPR+, i.e. already maximal..

In (5) the analysis of the noun phrase *hende den tossede malerinde* (her the crazy painter) is given.



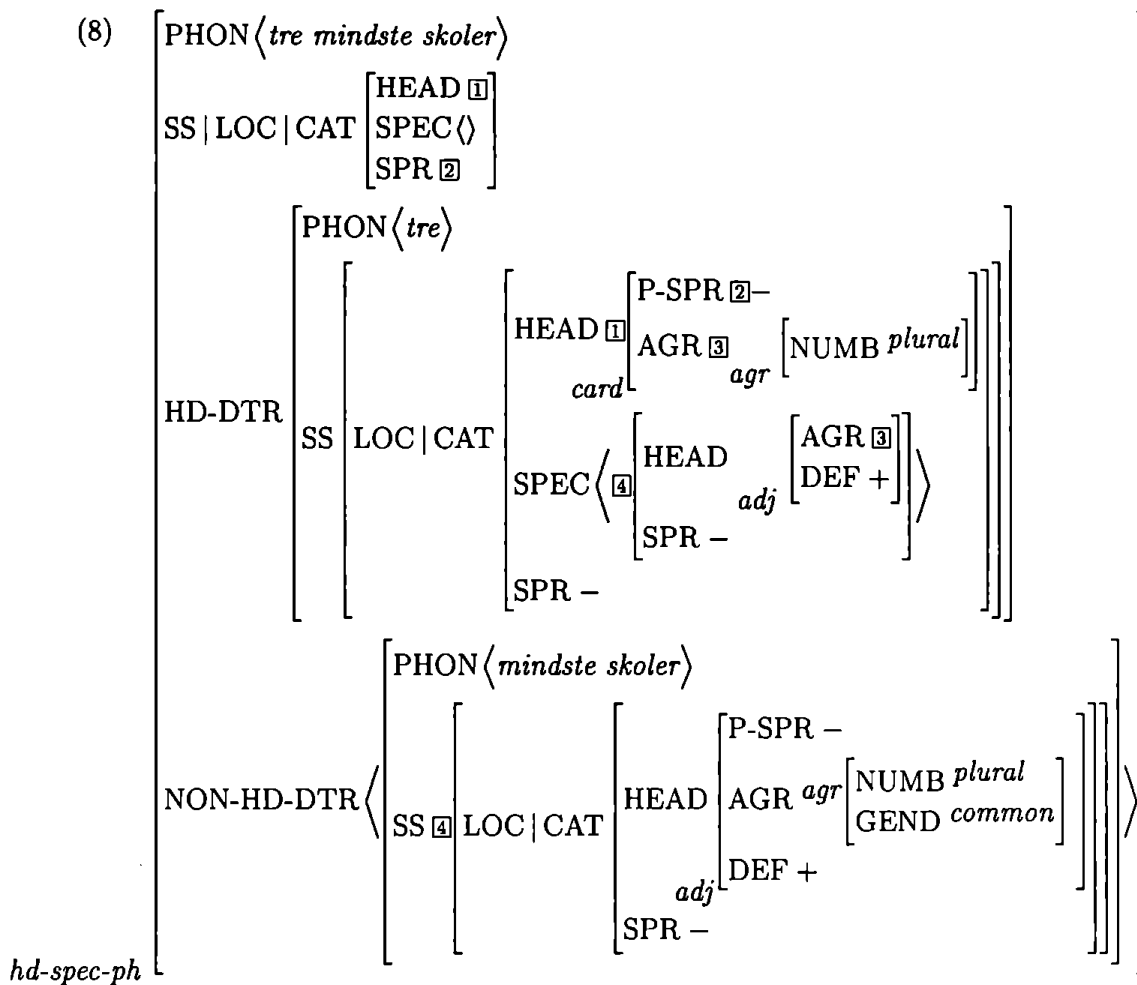
(5) shows that the noun phrase is indeed a maximal well-formed noun phrase. The selection restriction has been resolved by the unification of the underspecified selection constraint and the actual occurring constituent headed by a definite article, *defart*. The value of the specifier head's P-SPR attribute is structure shared with the SPR value of the projected phrase, the latter indicating whether the phrase is a maximal noun phrase, which it is in this case as it is SPR+.

In (6) and (7) the lexical entries for the cardinal specifier *tre* (three) are shown.



There are two entries for *tre* (three), and only (6) projects a maximal noun phrase because it is P-SPR+. The selection restriction for this version is something which is not *pronominal* and DEF-. The other version, (7), does not project a maximal phrase. The constraint on its selectee is also something which is not *pronominal*, however, there is no constraint on the value of DEF.

In (8) the analysis of *tre mindste skoler* (three smallest schools) is shown.



(8) shows that this phrase is not a maximal well-formed noun phrase. The selection restriction has been resolved by the unification of the underspecified selection constraint and the non-head constituent headed by an adjective, i.e. *adj*. The value of the specifier head's P-SPR attribute is again structure shared with the SPR value of the projected phrase. The projection is SPR- and consequently not a maximal noun phrase. What is important to note is that the version of the cardinal which projects maximal phrases cannot be used here as it constrains its selected constituent to be DEF-, but *mindste skoler* (smallest schools) is headed by a DEF+ adjective.

The two analyses show how specifier heads, the hierarchy and the definiteness feature are exploited to achieve correct analyses of the data.

4 Test with LKB

The proposed analysis of Danish noun phrases has been implemented in the LKB system which is a grammar and lexicon development environment for use with constraint-based formalisms (Copestake 1999). The system has been extensively tested with grammars based on HPSG theory. To test the analysis, a test suite

consisting of the data in section 1 has been parsed with the implemented grammar. The results are shown in (9)

- (9) 1 denne min begejstring 1
 2 denne den sidste Rabbitbog 1
 3 jeres den gamle grammofon 1
 4 deres den nærmeste nabo 1
 5 hende den tossede malerinde 2
 6 ham den forsvundne dreng 2
 7 de tre mindste skoler 2
 8 * tre mindste skoler 0
 9 de mange smukke ting 2
 10 *de mange smukkeste ting 0
 ;; Total CPU time: 1240 msec

The figure after each phrase gives the number of parses found. Examples 5, 6 and 9 get two parses. This is because the definite article and demonstrative pronoun have identical surface forms. Importantly the grammar correctly rules out 8 and 10.

5 Conclusion

In this article it has been pointed out that an important aspect of noun phrase analysis is the specification of selection restrictions for noun phrase specifiers in order to account for the combination of multiple specifiers. Danish data was presented which showed that definiteness plays an important role in this respect. This was reflected by the noun phrases containing multiple definite specifiers, and it was further shown that definiteness also plays a role among numerals and adjectives. A number of previous HPSG noun phrase accounts were discussed, and it was noted that they point towards two important properties of noun phrase structure. Firstly, specifiers must mark their projections. Secondly, a detailed classification of specifiers is required. An analysis was then presented which incorporated these properties together with a cross categorial definiteness feature to account for Danish noun phrase structure. The analysis has been implemented in the LKB system, and the results of parsing the Danish data were included, showing that the implementation, hence, the analysis, indeed provides an adequate analysis of the presented Danish noun phrases.

Footnotes

¹Hansen (1994) provides a description of the semantics of these constructions.

²In Pollard & Sag (1994) specifiers are categorized as functionals, i.e. words the semantic content of which "is purely logical in nature (perhaps even vacuous)" (Pollard & Sag 1994:45). In Pollard & Sag (1994:344-393), however, it is pointed out that many specifiers have semantic content and may take their own complements and specifiers, giving rise to complex specifier phrases which do in fact contain several specifiers. It is not quite clear whether a re-classification of such specifiers as non-functionals is intended in which case the SPEC attribute would be appropriate for

both functional and substantive categories. However, what is important is that no examples are provided of complex specifier phrases containing specifiers like articles or demonstratives, which means that multiple specifier sequences containing these are not accounted for. Even if these categories were contained in complex specifier phrases, the unaddressed problem consisting in defining head-dependent relations remains. Allegranza (1998) addresses this issue, and concludes that the establishment of such relations would be "quite arbitrary".

References

- Abney, S. 1987. *The English Noun Phrase in its Sentential Aspect*. MA. Thesis, MIT.
- Allegranza, V. 1998. Determiners as Functors: NP Structure in Italian. In S. Balari & L. Dini (eds) *Romance in HPSG*, CSLI Lecture Notes, Vol. 75. CSLI Publications: Stanford, 55-108.
- Börjars, K. 1994. *Feature Distribution in Swedish Noun Phrases*. Phd. dissertation, Department of Linguistics, University of Manchester.
- Chomsky, N. 1970. Remarks on Nominalization. In R. A. Jacobs, and P. S. Rosenbaum (eds) *Readings in English Transformational Grammar*. Waltham, MA: Ginn, 184-221.
- Copestake, A. 1999. *The (new) LKB system*. Stanford University.
- Delsing, L. 1993. *The Internal Structure of Noun Phrases in the Scandinavian Languages*. Malmö: Team Offset.
- Hansen, E. 1994. Kvalificeret bestemthed. Unpublished manuscript, Københavns Universitet.
- Jackendoff, R. 1977. *\bar{X} Syntax: A Study of Phrase Structure*. Cambridge, Massachusetts: MIT Press.
- Kathol, A. 1998. Determiners as Nominal Heads. In Gosse Bouma, Geert-Jan Kruijff & Richard Oehrle(eds) *Proceedings of FHCG'98*, Saarbrücken 1998, 136-143.
- Kolliakou, D. 1995. *Definites and Possessives in Modern Greek: an HPSG Syntax for Noun Phrases*. Ph.D. thesis, University of Edinburgh.
- Netter, K. Towards a Theory of Functional Heads: German Nominal Phrases. In J. Nerbonne, K. Netter and C. Pollard (eds) *German in Head-Driven Phrase Structure Grammar*. Stanford: CSLI.
- Pollard, C. & Sag, I. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.