A Phase-Based Approach to Scandinavian Definiteness Marking

Fabian Heck, Gereon Müller, and Jochen Trommer
University of Leipzig

1. Background

The present paper proposes a syntactic account of definiteness marking in Danish and Swedish that is based on the independently motivated notion of a phase in the sense of Chomsky (2001). The main claim is that prenominal definiteness marking in Scandinavian is the result of movement of a [def]-feature to the edge of a DP-phase in order to remove [def] from the impenetrable domain of the phase. The discussion here is restricted to standard Swedish and Danish, following Embick and Noyer (2001) and Hankamer and Mikkelsen (2005). For relevant discussion of other languages and/or dialects, see Delsing (1988, 1993), Julien (2005), Anderssen (2006), Alexiadou (2007), among many others (cf., in particular, Kramer (2007), who independently proposes an analysis of definiteness marking in Amharic that is very close to the present one in many respects.)

The core facts of Swedish and Danish definiteness marking are the following. To begin with, definiteness on unmodified nouns in Swedish and Danish is marked by a suffix (see (1)). However, in the context of a prenominal adjective a prenominal [def]-marker becomes obligatory (see (2-b) and (2-f)); Swedish (in contrast to Danish) also retains the suffix (see (2-b) vs. (2-c) and (2-e) vs. (2-f)).

(1) Swedish; simple noun
   a. mus-en
   b. * den mus
   c. hest-en
   d. * den hest

(2) Swedish; adjective plus noun
   a. * gamla mus-en
   b. den gamla mus-en
   c. * den gamla mus
   d. * gamle hest-en
   e. * den gamle hest-en
   f. den gamle hest

In what follows, we shortly outline two recent analyses of definiteness marking in Swedish and Danish, Embick and Noyer (2001) on the one hand and Hankamer and Mikkelsen (2005) on the other. We then present an alternative account of definiteness marking in terms of [def]-feature movement, arguing that it is able to account straightforwardly for all the observations made by Hankamer and Mikkelsen (2005). The paper concludes by illustrating that the present account is also able to account for interpretational issues of definiteness marking (that came up only recently in the literature), which the theories of Embick and Noyer (2001) and Hankamer and Mikkelsen (2005) cannot embrace for principled reasons.

* We would like to thank Hans-Martin Gärtner, Line Mikkelsen, and Sten Vikner for helpful comments.
2. Two previous analyses

2.1. A morphological account: Embick and Noyer (2001)

According to Embick and Noyer (2001) the feature [def] is a lexical property of D, sometimes spelled out as -en. They assume that N-to-D movement in Scandinavian (as in (4-a); see Delsing (1993)) is the default that applies in order to satisfy the constraints in (3-a,b). Furthermore, only if an A blocks N-to-D in the syntax as in (4-b) due to the Head Movement Constraint (HMC, Travis (1984)), the following postsyntactic morphological repair operations become active: First, -en on D is supported by an expletive d-host (see \( \odot \) in (4-c)); second, [def] on D is marked on N by insertion of a “dissociated” morpheme -en (see \( \circ \) in (4-c)). In contrast to Swedish, Danish lacks the second strategy (i.e., \( \odot \)) in this context.

(3) a. N-DEF: An N is marked as definite when D is [def]. b. D-DEF: A D that is [def] must have a host.

(4) a. DP b. DP c. DP

\[
\begin{align*}
\text{D} & \quad \text{NP} \\
\text{N} & \quad _2-\text{en} \\
\text{X} & \quad \text{N} \\
\text{D} & \quad \text{AP} \\
\text{A} & \quad \text{NP} \\
\end{align*}
\]

Under this analysis, d-support and insertion of a dissociated -en are last resort operations that do not apply unless N-to-D is blocked due to the presence of an intervening adjective. In other words, if nothing forces these last resort operations, N-to-D movement applies as the default.

The notions of last resort and default indicate a hidden competition in Embick and Noyer’s (2001) analysis. Moreover, condition N-DEF appears to be violated in (4-c) in Danish. The first point becomes even more apparent if Embick and Noyer’s (2001) theory is expressed in terms of optimality theoretic constraint interaction (Prince and Smolensky (2004)), which also makes understandable the violation of N-DEF in (4-c). An explicit reconstruction in these terms could look as follows. In addition to the constraints in (3-a,b), there are constraints against d-support (see (5-a), from Grimshaw (1997), proposed to control do-support in English), and against the insertion of dissociated morphemes (see (5-b)).

(5) a. *DISSOCIATION: No dissociated morphemes. b. FULL-INTERPRETATION: No d-support.

(6) a. Swedish ranking  
N-DEF, D-DEF, HMC \( \gg \) *DISSOCIATION, FULL-INT 
b. Danish ranking  
D-DEF, HMC \( \gg \) *DISSOCIATION, FULL-INT \( \gg \) N-DEF

The rankings for Swedish and Danish that derive the facts illustrated in (1) and (2) are given in (6).

2.2. A lexical account: Hankamer and Mikkelsen (2005)

The theory of Hankamer and Mikkelsen (2005) postulates a lexicon rule \( \mathbb{D} \) that converts an N into a definite D, which is spelled out as N-en (see (7)). Alternatively, they assume that a lexical definite D and an N can combine to form a DP in the syntax. Since As only merge with Ns, but not Ds, it follows that an N converted by \( \mathbb{D} \) cannot co-occur with an A. Thus the impossibility of A + N-en in Danish is derived.

\[
\begin{bmatrix} \text{PF} \quad \alpha \quad \text{CAT} \quad \text{N} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{PF} \quad \alpha + \text{en} \\
\text{CAT} \quad \text{D} \\
\text{DEF} + \end{bmatrix}
\]

Moreover, Hankamer and Mikkelsen (2005) assume that syntactically formed DPs are blocked by semantically equivalent DPs that are formed in the lexicon (see Sciullo and Williams (1987), Ackema and Neeleman (2004) on blocking). This prevents the prenominal marker to appear in contexts without A. Ns
that are modified by As cannot be blocked because there are no semantically equivalent nominals in the lexicon. To account for double marking in Swedish, Hankamer and Mikkelsen (2005) stipulate that the category change N—D by D is optional (making co-occurrence of A + N-en possible after all).¹

To summarize, just like the account of Embick and Noyer (2001), the theory of Hankamer and Mikkelsen (2005) inherently involves the concept of competition: The lexical strategy is preferred to the syntactic strategy if both are available in principle.

3. A phase-based account

3.1. Assumptions

The present phase-based approach is based on the idea that features (in particular the [def]-feature) can undergo movement (see Chomsky (1995)). Suppose (again following Chomsky (1995)) that movement in general, and feature movement in particular, involves copying, and that spell-out usually realizes the highest of a series of copies in a chain. Next, assume that alongside vP and CP, DP is a phase in the sense of Chomsky (2001) (see e.g., Svenonius (2004), Heck and Zimmermann (2004)). Phases, and thus DPs, are subject to the the Phase Impenetrability Condition (PIC, cf. Chomsky (2001)), a version of which is given in (8).

(8) PHASE IMPENETRABILITY CONDITION:
Material within a phase XP is not accessible to operations at ZP (the next phase) unless it is within the edge domain of XP.

(9) Edge domain:
The edge domain of a phase XP comprises the left area up to and including the leftmost overt element within XP.

A general consequence of the PIC is that any goal-feature that is supposed to value a probe-feature in a higher phase must first move to the edge domain of the current phase in order to be accessible for either Agree or further movement. This leads to the following well-known problem: If all movement is feature-driven due to LAST RESORT (LR, Chomsky (1995), Collins (1997); see (10)), then what triggers successive cyclic movement to the phase edge?

(10) LAST RESORT:
Movement must result in immediate feature valuation.

A standard answer given in the literature is that edge movement is driven by particular edge features, see Collins (1997), Fanselow and Mahajan (2000), Chomsky (2001), McCloskey (2002). This view is not unproblematic, though; for instance, it might be argued that the insertion of such edge features violates Chomsky’s (2001) INCLUSIVENESS Condition. In the present paper, we thus follow an alternative approach to the above mentioned problem. Namely, we follow Heck and Müller (2000, 2003) in assuming that movement to the phase edge is not feature-driven after all and therefore violates LR. However, such a violation is tolerated by the grammar if it helps to avoids a violation of the higher ranked constraint PHASE BALANCE (PB, see (11)).

(11) PHASE BALANCE:
For every probe in the numeration, a matching goal must be accessible in the current phase.

(12) Accessibility:
A matching goal is accessible iff it is in the edge domain of the current phase.

The definitions used here somewhat trivialize the complications taken into account in Heck and Müller (2000, 2003), but they suffice for the present purpose.² The core idea relevant for the present analysis

¹This raises the following question: If the category change does not take place and A merges with a so formed definite NP, then why does the prenominal marker enter the stage. Hankamer and Mikkelsen (2005) assume that this is the case because such definite NPs select a prenominal marker.

²For instance, if a probe in the numeration is matched by two goals, one in the numeration and one in the current phrase marker, then the probe must be able to satisfy PB without triggering movement of the goal in the current phrase marker (because the probe counts on being satisfied by the goal in the numeration at some later point).
is as follows. Assume that the feature [def] is a property of N in Scandinavian (cf. Delsing (1993); see also Borer (1999) on Hebrew). Next, suppose that [def] on an object must be accessible to a probe on v (the same holding for [def] on a subject and a probe on T), which is overtly signaled by verbal [def]-agreement in some languages (see Trommer (1995), Kiss (2002) on Hungarian). Now, if DP is a phase, then it follows that [def] must be in the edge domain of DP in order to be accessible for [def]-agreement. As we will see, this sometimes involves creating a higher copy of [def] by feature-movement, which is then spelled out as a prenominal [def]-marker.

3.2. Analysis

3.2.1. The core facts

In the absence of a prenominal A, N-[def] is the leftmost overt element within DP, and thus [def] is accessible. As a consequence, PB does not force [def]-movement, which is therefore blocked by LR. Since there is only one copy of [def] (namely on N), it trivially constitutes the highest one and is thus spelled out (namely as -en), see ① in (13-a). This holds for both Danish and Swedish.

In the presence of a prenominal A, PB forces movement of [def] from N across A into the edge domain of DP (in violation of LR) in order to remain accessible, see ① in (13-b). This creates two copies of [def], the higher copy of which is spelled out as an allomorph den/det/de, see ② in (13-b). Again, this is the case for both Danish and Swedish. However, for Swedish we need the additional stipulation that the lower copy of [def] is spelled out, too.

Note in passing that the present analysis does not depend on whether AP is the complement of D (see Abney (1987), Embick and Noyer (2001)) or whether it is in SpecN (see Svenonius (1992), Hankamer and Mikkelsen (2005)).

3.2.2. -ende-Nouns

Hankamer and Mikkelsen (2005, 97f.) observe that deverbal “common gender” nouns that are formed with the suffix -ende require the prenominal [def]-marker and are incompatible with the suffix -en in both Danish (see (14-a,b)) and Swedish. In contrast, neuter Ns formed with -ende are well-behaved in that they lack the prenominal marker in the absence of an A, see (14-c,d). Hankamer and Mikkelsen (2005) take this as an argument in favor of their lexical treatment, the idea being that a lexical rule can be sensitive to morphological properties while (post-)syntactic movement cannot.

Note that studerende formally looks like a present participle. Provided that it is treated as a participle in the syntax, i.e., that it modifies an empty N as in (15-a), it follows that its derivation involves movement of [def] from the empty N across the participle, resulting in a prenominal [def]-marker. Neuter -ende-forms on the other hand could be assumed to be real Ns; thus they do not contain any (overt) modifier, see (15-b). Movement of [def] is blocked by LR and no prenominal [def]-marker is spelled out.

\[
\begin{aligned}
3.2.1. & \text{The core facts} \\
3.2.2. & \text{-ende-Nouns} \\
(13) & \text{a. DP} & \text{b. DP} \\
& D & N \\
& \text{NP} & \text{NP} \\
& \text{[def]} & \text{[def]} \\
& \text{-en} & \text{den} \\
& ① & ② \\
(14) & \begin{array}{ll}
\text{Common gender -ende-noun} & \text{Neuter -ende-noun} \\
\text{a. * studerende-en} & \text{c. udseend-et} \\
\text{student-def} & \text{appearance-def} \\
\text{b. den studerende} & \text{d. * det udseende} \\
\text{def student} & \text{def appearance}
\end{array}
\]

Note that studerende formally looks like a present participle. Provided that it is treated as a participle in the syntax, i.e., that it modifies an empty N as in (15-a), it follows that its derivation involves movement of [def] from the empty N across the participle, resulting in a prenominal [def]-marker. Neuter -ende-forms on the other hand could be assumed to be real Ns; thus they do not contain any (overt) modifier, see (15-b). Movement of [def] is blocked by LR and no prenominal [def]-marker is spelled out.

\[
\begin{aligned}
\text{a. Deverbal common gender noun} \\
[\text{DP den [NP [AP studerende] [N Ø]]}]
\end{aligned}
\]
b. **Deverbal neuter noun**  
\[
\text{[DP D [NP anliggend-et]]}
\]

A welcome consequence of this analysis is that it automatically explains why common gender -ende-Ns in Swedish never bear a [def]-suffix, just as in Danish (see Hankamer and Mikkelsen (2005, 102)), in contrast to what is usually the case in Swedish. The idea is that spell-out of the lower copy is not possible in these cases, because the empty N is not an appropriate host to support the -en Suffix.

We hasten to add that Hankamer and Mikkelsen (2002) give arguments against the analysis in (15-a), claiming that forms like *studerende* should be treated as Ns in the syntax. We suspect that the arguments can be challenged.\(^3\) But should they turn out to prevail, then the present approach must resort to another explanation for the behavior of deverbal -ende nouns.

### 3.2.3. Relative clauses

The following observations of Hankamer and Mikkelsen (2005, 107f.) concern relative clauses. Restrictive relatives occur with prenominal or post-nominal [def]-marking in Danish and Swedish (see (16-a-c) and (16-d,e)). In contrast, appositive relatives always require post-nominal [def]-marking (see (16-a) and (16-d)). Finally note that the double marking in Swedish becomes optional in the context of restrictive relative clauses (see (16-b) and (16-c)).

(16)  

<table>
<thead>
<tr>
<th></th>
<th><strong>Swedish; restrictive and appositive</strong></th>
<th><strong>Danish; restrictive and appositive</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>mus-en som vi såg mouse-def that we saw</td>
<td>horse-def that won race</td>
</tr>
<tr>
<td></td>
<td>Swedish; restrictive only</td>
<td>Danish; restrictive only</td>
</tr>
<tr>
<td>b.</td>
<td>den mus som vi såg def mouse that we saw</td>
<td>def horse that won race</td>
</tr>
<tr>
<td></td>
<td>Swedish; restrictive only</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>den mus-en som vi såg def mouse-def that we saw</td>
<td></td>
</tr>
</tbody>
</table>

To derive these facts in the present phase-based theory we make the following assumptions. First, suppose that restrictive relatives are merged pre-nominally as the sister of N (see Smith (1964), Stockwell et al. (1973), where this is proposed, albeit for another reason; cf. also Fanselow (1986)). Restrictive relative clauses then obligatorily extrapose to the right (but are interpreted in-situ), an operation that we assume to be triggered by a feature on D (thus respecting LR). Further assume that appositive relatives are merged to the right of DP (or that they are not a proper part of their syntactic environment in narrow syntax to begin with, see, e.g., Emonds (1979), Cinque (1982)).

\(^3\)For instance, Hankamer and Mikkelsen (2002) observe that *studerende* cannot be modified by an adverbial as *ivrigt* ("eagerly"), while bona fide prenominal participles can (see (i-a,b)), suggesting that it cannot be a modifier.

(i)  

<table>
<thead>
<tr>
<th></th>
<th><strong>a. * den ivrigt studerende</strong></th>
<th><strong>b. den ivrigt læsende befolkning</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>def eagerly studying</td>
<td>def eagerly reading population</td>
</tr>
</tbody>
</table>

Actually, it turns out that (i-a) is grammatical, but only under an interpretation where the individual that the definite expression refers to is studying at the moment of utterance (i.e., (i-a) means “the eagerly studying person” but cannot mean “the eager student”). In a similar vein, Hankamer and Mikkelsen (2002) observe that (ii) can be true in a situation where the person referred to by *en studerende* is not studying.

(ii)  

<table>
<thead>
<tr>
<th></th>
<th><strong>Jeg så en studerende på gangen</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I saw a student in hallway</td>
</tr>
</tbody>
</table>

To be precise, (ii) is ambiguous between the two relevant readings, while the presence of the adverb disambiguates (ii-a). The ambiguity could be traced back to two different structures, one with *studerende* as an N, the other with *studerende* as a prenominal participle. However, an alternative interpretation would be to say that the participle *studerende* can occupy two different positions of modification within the DP, a higher or a lower one, resulting in different interpretations (see Cinque (1994)). The adverb in (i-a) would then be compatible with only one of these.
Finally, suppose that the phase head of the nominal domain is actually not D but n, a functional head between N and D (in the sense of Adger (2003)). As a consequence, PB will be relevant on the nP-level already, while the probe that triggers extraposition does not enter the structure before D is merged.

Under these assumptions, two options arise at the nP-level of a derivation involving a restrictive relative clause. Either PB-driven [def]-movement across the relative clause to the edge of nP applies, in order to render [def] accessible within nP. As before, such movement creates a higher [def]-copy spelled out as a prenominal marker. On the DP-level, feature-driven extraposition applies. Alternatively, PB can also be satisfied on the nP-level by early extraposition of the relative clause to the right of nP. Note that once extraposition has applied on the nP-level, [def]-movement to n is blocked by LR. Thus no higher [def]-copy is created. On the DP-level, string vacuous feature-driven extraposition applies. It follows that in the context of a restrictive relative either a prenominal or a post-nominal [def]-marker shows up. If appositive relatives are merged to the right of DP, they can never force [def]-raising to the edge of nP. Thus appositives only co-occur with a post-nominal [def]-marker. We must leave open here why optional deletion of the lower [def]-feature is possible in Swedish restrictive relatives (see (16-b,c)).

3.2.4. Post-nominal PPs

Finally, Hankamer and Mikkelsen (2005, 111f.) observe that post-nominal PPs never license the prenominal [def]-marker, but require the [def]-sufx. This holds for adjuncts as well as for complement PPs (see (18-a,b) and (18-c,d), respectively, for Danish).

(18) Post-nominal adjunct PP Post-nominal complement PP
a. gris-en med blå pletter c. forfatter-en til bog-en
pig-def with blue spots author-def to book-def
b. * den gris med blå pletter d. * den forfatter til bog-en
def pig with blue spots def author to book-def

These facts follow without further ado from the present theory if both types of PPs are merged post-nominally. In this position, they will never render [def] inaccessible, and thus LR blocks [def]-movement. Hence, there is only one low copy, spelled out as -en.

3.2.5. Interpretational issues

It has been claimed more recently (see Julien (2005), Anderssen (2006)) that the prenominal and the post-nominal [def]-marker are interpreted differently: Prenominal [def] denotes uniqueness, post-nominal [def] denotes specicity. Provided that this is correct, then a theory of definiteness marking should at least in principle be able to account for it. However, as we see it, the theories of both Embick and Noyer (2001) and Hankamer and Mikkelsen (2005) lack this capability for principles reasons. Here is why.

In the theory of Embick and Noyer (2001) the dissociated [def]-feature is copied post-syntactically, i.e., at a point of the derivation that is located behind the branching off to LF; therefore, dissociated [def]-features can not feed the LF-side of the derivation and consequently not contribute any meaning. For Hankamer and Mikkelsen (2005) it is crucial that blocking of a DP with prenominal [def]-marker by a lexically formed DP with post-nominal [def] is only possible if their interpretation is the same. This, of course, is incompatible with the idea that prenominal and post-nominal [def]-markers are interpreted differently.

For the phase-based approach proposed here no such principled problem arises: First, the [def]-feature is copied in the syntax and can thus feed LF; second, the competition involved in the present
analysis does not necessarily presuppose semantic equivalence of the competitors. Rather, we assume that the syntax simply provides structures, which are then accordingly interpreted. Depending then on its position in a structure, [def] is interpreted differently (see Diesing (1992), Chomsky (2001)).

4. Conclusion

To conclude, a syntactic (as opposed to a post-syntactic or a lexical) analysis of [def]-spread in Scandinavian is not only feasible, but attractive. Most of Hankamer and Mikkelsen’s (2005) observations fall out from the independently motivated movement to the edge of a phase, once the following assumptions are accepted: DPs (or nPs for that matter) are phases; [def] must value a probe on v (and T) and must therefore be part of DP’s edge domain; movement creates several copies, the higher of which (usually) undergoes spell-out.

References


