

Definite articles in Scandinavian: Competing grammaticalization processes in standard and non-standard varieties

Östen Dahl

Abstract

The standard Continental Scandinavian languages (Danish, Norwegian, and Swedish) are usually described as having two definite articles, one homophonous with the demonstrative pronoun *den* and preposed to the noun phrase (the P-article), the other manifested as an affix on the head noun (the S-article). The P-article is basically only used when the head noun is preceded by an attribute. There is however a difference between Danish and Swedish, with Norwegian in between, in that Danish suppresses the S-article whenever the P-article is used (*det store hus* ‘the big house’), whereas Swedish normally uses both articles: *det stora huset* ‘the big house’. The main claim of the paper is that the two articles are the result of two different processes of grammaticalization which have different geographical distributions, and that the variation we can see in the attributive constructions is the result of the competition between them about the same territory.

1. “Double determination”
2. The typological perspective: Special behaviour of definite NPs with attributes
3. Variation in the standard languages
4. Variation in spoken dialects in Sweden
5. Conclusions

1. “Double determination”

In the Scandinavian languages, two kinds of definite articles appear: preposed free articles (*den, det, de*), similar to the ones of West Germanic, and suffixed articles (*-en, -et* etc.). I shall refer to these as P-articles and S-articles, respectively. In addition, definiteness is also reflected by the choice of “weak” rather than “strong” endings of adjectives.

The story usually told about definiteness in Scandinavian is that there are two main patterns. Thus, in Danish (and marginally in Norwegian),

S-articles are used in definite noun phrases which do not contain a preposed modifier, and P-articles when the noun phrase contains an adjective or a quantifier.

	indefinite	definite
noun only	<i>et hus</i> 'a house'	<i>hus-et</i> 'the house'
adjective + noun	<i>et stor-t hus</i> 'a big house'	<i>det stor-e hus</i> 'the big house'

By contrast, in Swedish and most forms of Norwegian, while single nouns take an S-article, like in Danish, definite NPs with preposed attributes typically contain both a P-article and an S-article. The latter is a somewhat baffling phenomenon, which is variously referred to as “double determination”, “over-determination” (Norwegian *overbestemthet*) or “double articulation” (Plank 2003). It is illustrated here by Swedish examples:

	indefinite	definite
noun only	<i>ett hus</i> 'a house'	<i>hus-et</i> 'the house'
adjective + noun	<i>ett stor-t hus</i> 'a big house'	<i>det stor-a hus-et</i> 'the big house'

In this paper, I shall try to show how these major paradigms in the standard languages are just “the tip of the iceberg” in that they are part of a wider pattern of variation both within the standard languages themselves and within non-standard varieties, and that this pattern of variation can be seen as a manifestation of cross-linguistically attested grammaticalization paths.¹

2. The typological perspective: Special behaviour of definite NPs with attributes

In a language such as English, there is no difference in definiteness marking between noun phrases with and without adjectives and other attributes – in both cases one simply puts the definite article *the* in front of the noun phrase. As it turns out, English may be a minority case in this respect among languages that have some kind of definiteness marking. In many languages, complications of one sort or another arise in noun phrases

with modifiers. For simplicity, we shall here concentrate on definite nouns modified by adjectives.

2.1. Double (triple) definite marking

A relatively straightforward pattern is when the definite article is simply repeated, that is, both the noun and the adjective are preceded (or followed) by an identical morpheme. Modern Standard Arabic exemplifies this:

- (1) Standard Arabic
- a. *al-bayt*
DEF-house
'the house'
 - b. *al-bayt* *al-kabir*
DEF-house DEF-big
'the big house'

There are a great number of variations on this basic pattern. In many languages, adjectives may occur both before and after the noun. In those cases, double definite articles are more likely to appear when the adjective is postposed (Plank 2003: 342–347):

- (2) Modern Greek
- a. *i kondés fústes*
the short skirts
'the short skirts'
 - b. *i fústes i kondés*
the skirts the short
'the short skirts'
- (3) Yiddish
- a. *di grine oygn*
DEF green: PL eye: PL
'the green eyes'
 - b. *di oygn di grine*
DEF eye: PL DEF green: PL
'the green eyes'

In Old Icelandic, we may even find cases of double definite articles preposed to the adjective combined with a suffixed definite article on the noun. This triple marking is certainly a challenge for any theory that supposes that each morpheme fills a separate slot in the underlying structure. The following example is from the saga of Gísli Súrsson. The protagonist is having recurrent dreams, where two dream-women, one good and one bad, occur:

- (4) Old Icelandic
Hann segir, att nú kom at honom
 he say: PRS that now come: PRS to he: DAT
draumkona-n sú hin verri...
 dreamwoman-DEF DEF DEF worse
 ‘[Reporting Gísli’s answer to a question about his dreams:] He says that now came to him the evil dream-woman...’
 (Gísla saga Súrssonar 33)

2.2. Definite marking on adjectives only

This may also take different forms. In late Latin, demonstrative pronouns were used before postposed adjectives modifying proper names as in the following example:

- (5) (Late) Latin
Babylon illa magna
 Babylon DEM.F.SG great: F.SG
 ‘Babylon the Great’

In the Baltic and early stages of the Slavic languages, we find suffixed definite articles on adjectives, although the languages do not otherwise use definite marking:

- (6) Latvian
 a. *māja*
 house: NOM.SG
 ‘a/the house’

- b. *liel-a* *māja*
big-F.NOM.SG house
'a big house'
- c. *liel-ā* *māja*
big-F.NOM.SG.DEF house
'the big house'

Possibly, similar adjectival articles are the source of the “weak” endings on adjectives in Germanic, but there is no clear evidence on this point.

The rules for choosing between strong and weak endings differ between the Germanic languages, but it is still often the case that definiteness is involved, or at least that indefinite and definite noun phrases come out differently in some way, as in German:

- (7) German
 - a. *ein* *groß-es* *Haus*
one big-N.SG.NOM.STR house
'a big house'
 - b. *das* *groß-e* *Haus*
one big-N.SG.NOM.WK house
'the big house'

2.3. Articles cliticized to first element of NP

In another pattern, there is just one definiteness marking in each NP, showing up on the first element, whether it is a noun or an adjective, as in Amharic:

- (8) Amharic
 - a. *bet-u*
house-DEF
'the house'
 - b. *təlləq-u* *bet*
big-DEF house
'the big house'

What are the generalizations that can be made from the facts above? One observation is that NPs with adjectival modifiers tend to have at least as

much and often more definiteness marking than simple NPs. Furthermore, there is a clear tendency for any definiteness marking, excessive or otherwise, to show up next to or on the adjective. The existence of articles that mark adjectives only, as in Latvian or Old Slavonic, indicates that the initial stages of the grammaticalization of definite articles may be restricted to noun phrases containing modifiers. As a possible explanation of such a development, consider the fact that an adjective or relative clause (used restrictively) commonly singles out a subset within the set denoted by the head noun, contrasting it to its complement set. For instance, if I say *The rich countries won't suffer from the crisis* I contrast the countries that are rich against the countries that are not. Note that in English, there is a specific use of demonstratives to signal that a relative clause with this contrasting function follows:

(9) *those students who did not sign up for the exam*

Similar uses are plausible candidates for being the first step in the development of specific attributive articles. It is possible that they could later be extended into general definite articles, but as far as I know no such development has been properly documented.

3. Variation in the standard languages

In the introduction, I presented the standard view of double determination: Swedish has double definite articles, Danish doesn't and Norwegian is in between. The real picture is considerably more complex, however. There are many more possibilities to choose from than one would think and the variation within each language is considerable. If we take into account also non-standard varieties, the complexities increase further.

As we saw above, there are at least three possible definiteness markers in a Scandinavian NP with a preposed modifier: the preposed definite article – the P-article, the weak ending of the adjective, and the suffixed article on the noun – the S-article. In principle, each of these may be present or absent. This yields eight logical possibilities, which I shall denote as follows: $P\pm W\pm S\pm$, where P, W, and S stand for P-article, weak ending, and S-article, respectively, and + and – denote presence and absence of the marking.

In addition, the adjective may or may not be incorporated – that is, form a word with the following noun. In principle, this would yield another eight possibilities, but since incorporation seems incompatible with the P-article, we can denote the four remaining alternatives as IncW±S±. In actual practice, combinations of P+ and W– also seem excluded, which means that we have ten remaining possibilities.

We may identify two main domains within which there is considerable variation also in the standard languages, namely with what I shall call “selectors” and in “name-like uses”.

3.1. Selectors

This term is a cover term for three categories that are usually treated separately in grammars (all examples are Swedish):

1. a subset of what Teleman *et al.* (1999a: 435) call “relational pronouns”, namely *samma* ‘same’, “ordinative pronouns”, e.g. *först(a)* ‘first’, *sist(a)* ‘last’, *nästa* ‘next’, *förra* ‘previous’, “perspectival pronouns”, i.e. *höger/högra* ‘right (hand)’, *vänster/vänstra* ‘left (hand)’, *norra* ‘north’ etc., *övre* ‘upper’ and similar, *rätt* ‘right’, *fel* ‘wrong’, and *ena* ‘one (of)’
2. ordinal numerals
3. superlatives

All these categories share a common semantics – they are all “inherently definite” in that the noun phrases they are used in normally have definite reference by virtue of their meaning. We may note that Teleman *et al.* treat *först(a)* both as an ordinative pronoun and an ordinal numeral. In addition, one could claim that it is the superlative form of a defective adjective (with no basic form).

In English and German, we find certain examples of noun phrases with selectors but without definite articles, especially with deictic uses of phrases such as *next year* or *nächstes Jahr*. In the Continental Scandinavian the tendency for selectors to disfavour definiteness marking is more general, applying to all the three standard languages² and to all the sub-categories listed above, but with a quite astonishing degree of variation. In the table below, I have chosen three frequent collocations from each sub-category, ‘the right hand’, ‘the first time’, and ‘the eldest son’, indicating

(on the basis of an Internet³ search) what possibilities are possible. The forms in bold are the preferred choices in each language. Forms in parentheses appear to be marginal or archaic. The grammaticality judgments should all pertain to uses with definite reference.

Table 1. Variation in definiteness marking with selectors in the Continental Scandinavian standard languages

	Danish	Norwegian	Swedish
P–W–S–	<i>*først gang</i> <i>*ældst søn</i>	<i>*først gang</i> <i>*eldst sønn</i>	<i>höger hand</i> <i>*først gång</i> <i>*äldst son</i>
P–W+ S–	<i>højre hånd</i> <i>første gang</i>	<i>høyre hånd</i> <i>første gang</i>	<i>*högra hand</i> <i>första gång</i>
P–W+S+	<i>*højre hånden</i> <i>*første gangen</i> <i>*ældste sønnen</i>	<i>høyre hånden</i> <i>første gangen</i> <i>eldste sønnen</i>	<i>högra handen</i> <i>första gången</i> <i>äldsta/äldste</i> <i>sonen</i>
P+W+S–	<i>den højre hånd</i> <i>den første gang</i> <i>den ældste søn</i>	<i>(den høyre hånd)</i> <i>den første gang</i> <i>den eldste sønn</i>	<i>*den högra hand</i> <i>*den första gång</i> <i>*den äldste son</i>
P+W+S+	<i>*den højre hånden</i> <i>*den første gangen</i> <i>*den ældste sønnen</i>	<i>den høyre hånden</i> <i>den første gangen</i> <i>den eldste sønnen</i>	<i>den högra handen</i> <i>den första gången</i> <i>den äldsta/e sonen</i>
IncW–S+	<i>højrehånden</i> <i>*førstgangen</i> <i>*ældstsønnen</i>	<i>høyrehånden</i> <i>*førstgangen</i> <i>eldstsønnen</i>	<i>högerhanden</i> <i>(førstgången)</i> <i>*äldstsonen</i>

It should be noted that the word *højre/høyre* ‘right’ in Danish and Norwegian does not distinguish between a form with a zero ending and a form with a weak ending. The corresponding possibilities are therefore collapsed in those languages.

In all, six different patterns are represented. The average number of grammatical possibilities (including marginal uses) for each collocation is 3.1. In Danish, the average is 2, in Norwegian 4, and in Swedish 3.3.

We see that there are some clear constraints that reduce the range of the variation. In Danish, not unexpectedly, no combinations with both P-articles and S-articles are allowed. Conversely, in Swedish, a P-article cannot be used without an S-article. On the whole, wholly unmarked NPs (P–W–S–) are not allowed, except Swedish *höger hand*. The most liberal

language is Norwegian. It differs from Danish in allowing double determination and indeed preferring the pattern P+W+S+ for ‘the eldest son’. It thus differs also from Swedish, where the preferred choice is P–W+S+.

We may summarize the tendencies by looking at the number of markings in the preferred choices for each language, as in the following table:

	Danish	Norwegian	Swedish
P-article	1	1	0
S-article	0	1	2

In Swedish, preferences vary greatly within the types. With superlatives, the P–W+S+ is, as we saw, dominant in cases like ‘the eldest son’. This pattern may be restricted to kin terms preceded by an adjective indicating a ranking in age (such as *äldst* ‘eldest’, *youngst* ‘youngest’, *störst* ‘biggest’, *minst* ‘smallest’). Even the same adjectives with non-kin nouns are quite different. I looked for (*den*) *äldste medlemmen* ‘the oldest member’ on the Internet and found 36 occurrences with and only one without the P-article. To complicate things further, the pattern P+W+S– does occur marginally with superlatives in Swedish, e.g. *med den största försiktighet* ‘with the greatest care’, although the most common pattern here is P–W+S–, *med största försiktighet*. This appears to be a feature of the written language.

3.2. Name-like uses

A noun phrase containing an adjectival modifier and a head noun may be lexicalized as a proper name, as in *the White House*. In this English phrase, the proper name status is reflected in the stress pattern, which is the one normally used with compounds. In Scandinavian, we can see a tendency towards decreased definiteness marking. There is great variation in how this tendency is realized, however. Starting with ‘the White House’, the most common translations into the standard languages⁴ are as follows:

- Danish: *det Hvide Hus* (P+W+S–)
- Norwegian Bokmål: *det Hvite hus* (P+W+S–)
- Norwegian Nynorsk: *det Kvite huset* (P+W+S+)
- Swedish: *Vita huset* (P–W+S+)

Comparing this to the “normal” pattern in each language, we can see whereas Danish and Nynorsk do not make any changes here, Bokmål and Swedish drop one article each – but in Bokmål it is the P-article that survives, in Swedish it is the S-article.

There is, however, a further alternative, not always mentioned in grammars, viz. no marking at all (P–W–S–), as in *Svensk uppslagsbok* ‘The Swedish Encyclopedia’. This pattern seems to be more common in Danish and Norwegian than in Swedish, although it is found there, too, as the previous example shows. Thus, the Nordic Council is called *Nordiska rådet* (P–W+S+) in Swedish but *Nordisk råd* (P–W–S–) in the other languages. The two patterns (P+W+S–) and (P–W–S–) are sometimes used in competition with each other about the same entity. Thus, in Copenhagen I have observed street signs saying both *Den Gamle Kongevej* (P+W+S–) and *Gammel Kongevej* (P–W–S–) ‘Old King’s Road’. In addition, the pattern P–W+S–, that is, with the weak adjective ending as the sole definiteness marker, seems to have been possible in older Danish, as in the 17th century *Danske Lov* ‘the Danish Law’, and is preserved in names such as *Store Bælt* ‘the Big Belt’ (the sound between the islands of Sjælland and Fyn).

Adjective-noun compounding is a common strategy with proper names in all the languages, particularly in place-names. Even here there is variation, though. Most saliently, the compound name may appear with and without an S-article. It appears that Danish is more reluctant than Norwegian and Swedish to use the S-article here. Contrast Danish *Nygade* (IncW–S–) with Norwegian *Nygaten/Nygata* (IncW–S+) and Swedish *Nygatan* (IncW–S+). Interestingly, an Internet search for *Nygaden*, which would be the Danish definite form, yields no hits in Denmark but a couple of references to Norwegian 19th century usage, suggesting that this pattern was established already at the time when written Norwegian was otherwise essentially identical to Danish. On the other hand, there is also variation within the individual languages. In Swedish, a relatively systematic pattern is found, in that a name such as *Långsjön* ‘the Long Lake’ (IncW–S+) with the S-article, would refer to the lake itself whereas a village next to it would be called *Långsjö*, (IncW–S–) without definiteness marking.

So far, the cases we have been discussing have been clearly lexicalized proper name uses of adjective-noun combinations. However, the label “name-like uses” was chosen to include also other cases which are less clearly proprial, and where reduced definiteness marking may show up in Swedish. For instance, if you own two houses of different size next to each

other, it is very natural to call them *stora huset* ‘the big house’ and *lilla huset* ‘the little house’ (both P–W+S+), even before these denominations have become so “entrenched” that it is natural to use capital letters. Cf. the following example:

- (10) Standard Swedish
Mitt i byn ligger en gård med en stor kastanj på gräsmattan, två stugor, stora huset och "hönshuset" och ett stall.
 middle in village: DEF lie: PRS a farm with a big chestnut on lawn: DEF, two hut: PL big: WK house: DEF and hen_house: DEF and a stables
 ‘In the middle of the village there is a farm with a big chestnut tree on the lawn, two huts, the big house and the “hen house” and a stables.’ (Internet)

An alternative here is to use a compound. The following example, which is the continuation of the text in (10), illustrates both possibilities:

- (11) Standard Swedish
Här under ser du en bild på storhuset sett från gräsmattan, en bild på stora rummet och en på öppna spisen.
 here below see: PRS you a picture on big_house: DEF seen from lawn: DEF a picture on big: WK room: DEF and one on open: WK fire: DEF
 ‘Here below you see a picture of the big house (IncW+S+) from the lawn, a picture of the big room (P–W+S+) and the open fire-place (P–W+S+).’ (Internet)

Teleman *et al.* (1999b: 19) say that the P-article-less usage is possible only if the adjective has a restrictive interpretation and the referent can be identified in the speech situation or through previous experience “but not if it is identified anaphorically, through the [linguistic] context” (my translation). They give the following example in support of the claim:

- (12) Standard Swedish
Han hade två knivar med sig.
 he have: PST two knife: PL with REFL.3SG

<i>Han valde</i>	{ <i>den/*Ø</i> }	<i>stora</i>	<i>kniven</i>
he choose: PST	DEF	big: WK	knife: DEF
<i>att stycka</i>	<i>älgen</i>	<i>med.</i>	
to cut_up: INF	elk: DEF	with	

‘He had brought two knives. He chose the big knife to cut up the elk with.’

While the anaphoricity constraint seems to go in the right direction, the preceding examples illustrate that it is not so easy to define what it means to be identified anaphorically. If an object is referred to by a description that characterizes it as having a permanent role in a certain structure (e.g., being “the big house” on a certain farm or “the big room” in a house), it appears to be less important how it is introduced.

In view of the fact that the (P–W+S+) is not in general possible in conventionalized names in Danish and Norwegian, it is perhaps not so astonishing that the uses I have mentioned here are restricted to Swedish.

In both NPs with selectors and name-like NPs, we thus see a general tendency to have less definiteness marking than in other adjective-noun combinations. This tendency is manifested in rather different ways in the standard Scandinavian languages, however. In Swedish, the reduction hits primarily the use of the P-article, whereas in the other languages, it is either the S-article that is reduced, or both articles are reduced at the same time. It is thus tempting to assume that the two articles do not have the same strength in the different languages, the P-article being strongest in Danish and the S-article strongest in Swedish. This assumption is corroborated if we look at another domain, viz. the use of definiteness marking after determiners such as demonstrative pronouns.

At least three series of demonstrative pronouns commonly occur in the standard Scandinavian languages:

1. *den, det, de*
2. *denne/denna, dette/detta, disse/dessa*
3. *den här/där, det här/där, de här/där*

The first two series are found in all the languages, but the third only in Swedish. In Danish, a noun following a demonstrative never gets an S-article, in Norwegian, it always gets it. Swedish is more complicated. In the written standard language, nouns preceded by demonstratives in the first and the third series always take the S-article, but the second series does not.

In the spoken language, on the other hand, *denne* etc. can also combine with an S-article. It seems impossible to avoid talking about regional varieties here. The patterns that naturally occur in spoken language are (1) and (2) with S-articles in southern and western Sweden and (1) and (3), also with S-articles, in central Sweden. The articleless use of (2) is thus essentially a feature of the written language, and if it is disregarded, what we find is that the usage with demonstratives is consistent with the view that the S-article is weaker in Danish than in the other two languages.

Actually, given that the P-article is historically a demonstrative pronoun, and is indeed still formally indistinguishable from one, except for being prosodically less prominent, double definite articles may be seen as a special case of a more general phenomenon in which the S-article is used in the presence of a preposed determiner.

Summing up the use of the two articles in the standard Scandinavian languages, we can see that Danish uses the P-article most consistently and the S-article least often of the three languages, whereas Swedish has the strongest tendency to omit the P-article. We shall now zoom in on the variation found in non-standard varieties.

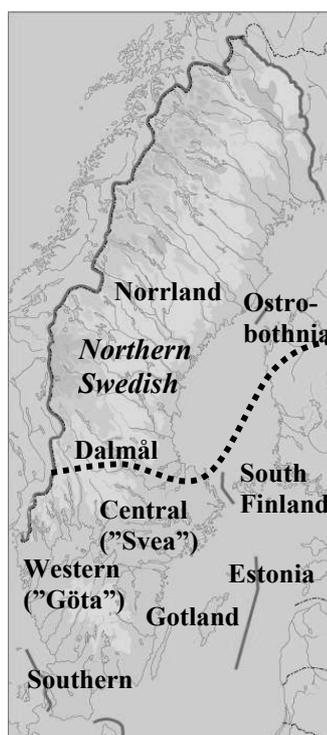
4. Variation in spoken dialects in Sweden

4.1. General

We now turn to the realization of definite NPs with preposed modifiers in what is traditionally called “Swedish dialects”. This notion is far from unproblematic. Basically, the Scandinavian varieties (excluding Icelandic and Faroese) constitute a dialect continuum, and any attempt to draw borderlines within them will meet with difficulties. The present-day political borders between the countries do not correspond well to the internal structure of the dialect continuum, and the diversity within the countries may be greater than the differences between the national standards. The dialects spoken within the present borders of Sweden thus do not constitute a “natural class” from the linguistic point of view. However, in comprising areas that have historically belonged to all three major Scandinavian-speaking countries, it does in fact give a cross section that may tell us more than if we were just looking at a well-defined branch of the Scandinavian languages.

In a way this is just post-hoc apologetics, since the material I will mainly rely on was not delimited by the aims of this investigation. I am

speaking of a parallel corpus of 35 dialects – henceforth referred to as “the Cat Corpus” – all spoken in Sweden except “Gammalsvenska”, the dialect spoken in Gammalsvenskby/Zmejevka in Ukraine:⁵ (In addition there is also a Standard Swedish version.) The texts are translations of a 6250 word story, “Granny’s Cat”, written by Rut “Puck” Olsson, who also arranged the translations, which are all by native speakers. (The spelling used in the examples below is in principle that of the original texts – hence the heterogeneity.) Although the coverage is somewhat patchy (more than a third of the translations are from the province of Dalarna), all major parts of Sweden are represented. In spite of the relative brevity of the texts, it is possible to see fairly clear patterns in the use of definiteness marking in adjective+noun combinations.



Map 1. Major “Swedish” dialect areas

Map 1 shows the major dialect areas within the political boundaries of Sweden and the Swedish-speaking areas in Finland and Estonia. The latter are usually regarded as one “Eastern Swedish” area but are different

enough to deserve being kept apart. The dialectal divisions in Sweden still reflect the political structure of the area in the Middle Ages.

The Southern area basically covers the originally Danish provinces in the south but extends into the province of Småland. The Western (“Göta”)⁶ and Central (“Svea”) areas correspond to the major medieval political division of the core Swedish provinces. The case of the province of Dalarna (Dalecarlia) is rather special. The vernaculars traditionally spoken in the northern part of the province – referred to as Dalecarlian in English, *dalmål* in Swedish, and *dalska* by the speakers themselves – are highly deviant with many conservative traits but also with innovations which make them incomprehensible to outsiders. The internal variation is also quite extreme, making comprehension across parish borders difficult. The southern part of the province of Dalarna is part of the Central Swedish mining district called Bergslagen, which could be considered a dialect area of its own, although I have not marked it on the map to preserve legibility. It has some common features with the Dalecarlian. Highly deviant and conservative vernaculars are also found in parts of Norrland, which is a peripheral area (although making up two thirds of Sweden), parts of which were colonized relatively late, with a significant Saami and Finnish-speaking population.

For the purposes of this investigation, it is convenient to separate out a “northern area”, comprising all the dialects in the Cat Corpus spoken in Dalarna and Norrland.

4.2. P-articles in the dialects

We shall start by considering the P-article in the Cat Corpus. This turns out to be somewhat more difficult than might be expected – it is not at all clear what should count as a P-article. The problem is that there are at least three different types of preposed determiners that might qualify as articles here. In addition to the usual one, manifested as a form of the series *den/det/de*, we also find cognates of the Standard Swedish 3rd person pronouns *han* and *hon*, and also combinations of *den*, *han* etc. with an adverb ‘there’. Thus, we find translations as the following:

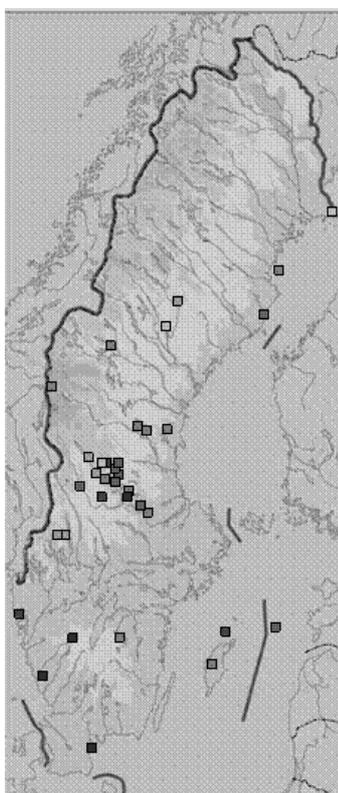
- (13) P-article types in Swedish dialects [Context: Granny starts talking into one end of the receiver of her new telephone, which the cat thinks looks like a sausage.]

- a. P-article Type I. *Den, det, de*.
Timmele (Västergötland)
Den andre kôrvännen hade ho så gött
DEF other: WK sausage_end: DEF have: PST she as well
sûm stôppat in i örat.
as stuff: SUP in in ear: DEF
'**The other sausage end** she had virtually stuffed into her ear.'
(Cat Corpus)
- b. P-article Type II. *Han, hon, häd* etc.
Ore (Dalarna)
Ann ar kôrwändn add o sô gött
DEF other: WK sausage_end: DEF have: PST she as well
summ stuppa in i ära.
as stuff: SUP in in ear: DEF
'**The other sausage end** she had virtually stuffed into her ear.'
(Cat Corpus)
- c. P-article Type III. *Den/han där* etc.
Orsa (Dalarna)
Dendâ ödör endân stuft -o in i ära.
DEF other end: DEF stuff: PST she in in ear: DEF
'**The other end** she stuffed into her ear' (Cat Corpus)

Table 1 shows the distribution of the three types in the Cat Corpus. What we can see is that Type I, the "standard" P-article, is found all over the country, but its use seems much more systematic in the southern part of Sweden than in the rest of the country, and in several parts of the northern area, it is used very little.

Types II and III are only found in the northern area and in the Gammalsvenskby dialect. Type II reflects differences in the pronominal systems: (i) in the northern area (also elsewhere: Finland, Estonia (incl. Gammalsvenskby), partly Gotland and the province of Uppland in the central area), *han* and *hon*, which in the standard languages are only used as freestanding third-person pronouns, can also be used as noun-modifying demonstratives; (i) in roughly the same areas, the standard language pronoun *det* 'it, that' (in all its various functions) corresponds to forms going back to something like **häd* (*he, hä, ä, äđ* etc.).

For Type III, the material is a bit shaky – the majority of the occurrences of Type III are found in one and the same sentence, and it cannot be excluded that the translators intend it as an ordinary demonstrative pronoun. However, it is quite clear that in some dialects, this type functions straightforwardly as a P-article. Cf. for instance (13b) from Orsa, where an interpretation ‘that other end’ makes little sense. This usage is also well-known from different parts of Swedish-speaking Finland (Anne-Marie Ivars, p.c.). In the Cat Corpus, the clear cases are restricted to the Dalecarlian area. In Orsa, this is in general the preferred pattern for definite adjective-noun combinations (Eva Olander, p.c.).



Map 2. Total distribution of P-articles in the Cat Corpus

Looking at the total distribution of P-articles (Map 2), we can see that it is quite uneven. Again, the highest numbers are all in southern and western

Sweden, and the northern area includes several dialects where there are hardly any occurrences at all in the corpus. In fact, it can be shown that in these dialects, there are alternative ways of expressing definite adjectives in NPs which do not involve P-articles. One is adjective incorporation, or as it is traditionally described, compound formation, as in (14)–(15):

- (14) Nås (Dalarna)
...enn ekkôrr sômm satt ti stortâll-n
 a squirrel that sat I big_fir-DEF
bôrrta brônn ô flast kâkklôr.
 beyond well: DEF and peel: DEF cone: PL
 ‘...a squirrel who sat in **the big fir-tree** behind the well, peeling cones’ (Cat Corpus)

- (15) Färila (Hälsingland)
Kammârfönstre va öppet
 chamber_window: DEF be: PST open: N
å vitgardin-ân fladdrâ.
 and white_curtain: DEF.PL flutter: PST
 ‘The bedroom window was open and **the white curtains** were fluttering’ (Cat Corpus)

In many dialects in the northern area, this is the normal way of forming definite adjective-noun combinations. Definiteness is then marked only once: by an S-article on the noun. It should be noticed that these combinations are pronounced with the main stress on the first element – compound stress – rather than on the second element – regular phrasal stress. However, it is somewhat difficult to tell exactly what is adjective incorporation and what is not. The problem is that in the same area where adjective incorporation occurs, there is a general process of elision of final vowels or apocope which hits the weak endings of adjectives. Sometimes there are combinations of endless adjectives and definite nouns in the corpus that are not written together as one word, and from the written form it is not possible to establish the intended prosody:

- (16) Leksand (Dalarna)
 ... *en ekorrâ, sôm sått ti stur tal'n*
 ... a squirrel that sat i big fir-DEF

borta *brona* *å* *skalå* *kôttår.*
 beyond well: DEF and peel: DEF cone: PL
 ‘...a squirrel who sat in **the big fir-tree** behind the well, peeling
 cones’ (Cat Corpus)

Notice, however, that in this example and several other similar ones, there is no P-article. Such occurrences of what looks like a P–W–S+ pattern, but which are rather P–W+S–, with the weak ending apocopated, are also attested from Finland and Estonia.⁷

It may be noted that adjective-noun compounds have a relatively high frequency also in Standard Swedish. Thus, the example (16) comes out as (17) in the Standard Swedish version:

(17) Standard Swedish
...en *ekorre* *som* *satt* *i* *stortallen*
 a squirrel that sat I big_fir-DEF
bortom *brunnen* *och* *skalade* *kottar.*
 beyond well: DEF and peel: DEF cone: PL
 ‘...a squirrel who sat in **the big fir-tree** behind the well, peeling
 cones’ (Cat Corpus)

Significantly, however, there are no examples from the southern or western area of this pattern.

When adjectives occur without a head noun, the incorporation strategy obviously cannot be used. In all the standard languages and in the southern, western and central dialect areas in Sweden, the common strategy is to use a P-article and a weak ending on the adjective – that is, what you get by simply omitting the noun from the normal definite adjective + noun construction. In the northern area, there is again a strategy not involving the P-article – what happens is that the adjective is equipped with an S-article instead. For the superlative ‘the youngest (son)’ this is attested for ten dialects in the northern area, cf. for instance (18):

(18) Mora (Dalarna)
Ynggst-n *wa* *jenn* *slåjk* “*waturallår*” ...
 young: SUPERL-DEF be: PST a such water-navvy
 ‘**The youngest one** was a hydro-electric plant construction worker’
 (Cat Corpus)

Adjectives with S-articles are in fact marginally acceptable also in Standard Swedish, in particular with examples such as *yngrsten* ‘the youngest one’. (They would normally be pronounced with an acute pitch accent which suggests that the S-article is added to the stem rather than to a form with a weak ending.)

The other occurrence of a headless adjective in the corpus is in the context ‘the first (thing) he saw when he came out into the yard’. Here, there is only one unambiguous case of an S-article:

- (19) Älvdalen (Dalarna):
Fuäst-ar *an* *såg,*
 first: DEF he see: PST
mes an kam auto gardn...
 when he come: PST out_on yard: DEF.ACC
 ‘**The first (thing)** he saw when he came out into the yard’ (Cat Corpus)

Many other dialects in the area use what seems to be a bare adjective with a normal weak ending. Since the neutral S-article *-t* would normally be reduced to zero (by a phonological process) here, it is not possible to judge what the status of these forms is, cf. for instance (20).

- (20) Nederkalix (Norrbotten)
Förste *'n* *så,*
 first: DEF he see: PST
når 'n kãm öut opa går'n...
 when he come: PST out on yard: DEF
 ‘**The first (thing)** he saw when he came out into the yard’ (Cat Corpus)

There are also many examples from the northern area of P-articles in this context:

- (21) Sävar (Västerbotten)
De förstü *han* *såg,*
 DEF first: DEF he see: PST

då han kömma uta går'n...
 when he come: PST out_on yard: DEF
 'The first (thing) he saw when he came out into the yard' (Cat Corpus)

It is possible that this is one of the first uses where P-articles show up when they spread to a new dialect. As with the use of P-articles in general, it is least common in the generally most conservative dialects (as those in the Dalecarlian area), suggesting that P-articles are in general a recent phenomenon in the northern area.

This takes us to a general summary of the use of P-articles in the northern area and the Swedish-speaking parts of Finland and Estonia: On one hand, the common Scandinavian P-article *den*, etc., is either absent or only weakly represented in these areas. At the same time, there is also evidence of wholly or partly separate grammaticalization processes, resulting in alternative forms of preposed definite articles.

We can thus see that in large parts of the Swedish dialect area, the P-article is significantly weaker than in Standard Swedish. We shall also see, however, that there are geographical areas where the P-article is stronger, that is, used more often. Consider NPs with selectors, discussed above in section 3. We saw that in the phrase 'the eldest son', P-articles were preferred in Danish (P+W+S-), and Norwegian (P+W+S+), while the pattern P-W+S+ was the most common in Standard Swedish. Most dialects in the Cat Corpus follow Standard Swedish in this regard, but four of them do use a P-article in the relevant sentence:

(22) Standard Swedish
Äldsta pojken hade
 old: SUPERL.WK boy: DEF have: PST
rest till Amerika.
 travel: SUP to America
 'The eldest boy (i.e. Granny's son) had gone to America' (Cat Corpus)

(23) Träslövsläge (Halland)
Den gamlaste päjken hade
 DEF old: SUPERL.WK boy: DEF have: PST

fåt *te* *Amerka.*
 travel: SUP to America
 (=22)

- (24) Sotenäs (Bohuslän)
Den gam! aste ***pôjken*** *hade*
 DEF old: SUPERL.WK boy: DEF have: PST
fore *te* *Amerka.*
 travel: SUP to America
 (=22)

- (25) Timmele (Västergötland)
Den äldste ***pâjken*** *hade*
 DEF old: SUPERL.WK boy: DEF have: PST
fôert *te* *Amerka.*
 travel: SUP to America
 (=22)

- (26) Österlen (Skåne)
Den älste ***pâgen*** *hadde*
 DEF old: SUPERL.WK boy: DEF have: PST
fâred *ti* *Amerika.*
 travel: SUP to America
 (=22)

It is hardly a coincidence that these are precisely the dialects spoken in the southern and western areas. The pattern here is consistently P+W+S+, that is, in this respect these dialects line up with Norwegian. In another case of an NP with a selector, viz. the phrase 'the first time she saw the cat', on the other hand, there is not a single occurrence of a P-article in the Cat Corpus. However, as we see in Table 2, this is a context where all the Scandinavian standard languages also lack P-articles. This means that the southern and western dialects in Sweden display the same pattern as Danish and Norwegian with respect to P-articles with selectors. They clearly deviate from Danish, however, in their more liberal use of S-articles in these contexts.

It is more difficult to document name-like uses of adjective-noun combinations in the southern area – there are no occurrences in the Cat Corpus. Traditional names of cultivated fields in Skåne (the southernmost

province) such as *Norre vång* ‘the northern openfield’ (Dahl 1989) suggest the Danish P–W+S– pattern, while modern names such as *Norra Fälåden* ‘the northern pasture’ (a popular name of suburbs in Scanian towns) exemplify the Swedish P–W+S+ pattern.

Table 2. Distribution of P-articles in the Cat Corpus

Location	Province	Larger area	<i>Den</i>	<i>Han</i>	<i>Den där/ han där</i>	Sum
Brösarp	Skåne	Southern	6	0	0	6
Tråslövslåge	Halland	Southern	6	0	0	6
Timmele	Våstergötland	Western	7	0	0	7
Lau	Gotland	Gotland	3	0	0	3
Fårö	Gotland	Gotland	4	0	0	4
Lysekil	Bohuslån	Western	5	0	0	5
Våstra Ryd	Östergötland	transitional	3	0	0	3
Våstra Åmtervik	Vårmland	transitional	1	0	1	2
Mangskog	Vårmland	transitional	1	0	1	2
Grytnås	Dalarna	Bergslagen	2	0	1	3
Stora Tuna	Dalarna	Bergslagen	3	2	0	5
Aspeboda	Dalarna	Bergslagen	3	0	0	3
Bingsjö	Dalarna	Bergslagen	3	0	1	4
Nås	Dalarna	Dalecarlian	2	3	0	5
Malung	Dalarna	Dalecarlian	2	2	0	4
Leksand	Dalarna	Dalecarlian	3	0	1	4
Sollerön	Dalarna	Dalecarlian	0	1	2	3
Boda	Dalarna	Dalecarlian	2	1	1	4
Råttvik	Dalarna	Dalecarlian	2	1	1	4
Mora	Dalarna	Dalecarlian	0	1	1	2
Ore	Dalarna	Dalecarlian	0	3	1	4
Orsa	Dalarna	Dalecarlian	2	0	3	5
Våmhus	Dalarna	Dalecarlian	0	0	1	1
Ålvdalen	Dalarna	Dalecarlian	1	0	1	2
Jårvsö	Hålsingland	Norrland	2	0	1	3
Forsa	Hålsingland	Norrland	3	0	0	3

Location	Province	Larger area	<i>Den</i>	<i>Han</i>	<i>Den där/ han där</i>	Sum
Färila	Hälsingland	Norrland	2	1	0	3
Ljusnedal	Härjedalen	Norrland	2	1	0	3
Lit	Jämtland	Norrland	2	1	0	3
Junsele	Ångermanland	Norrland	1	0	0	1
Sävar	Västerbotten	Norrland	4	0	0	4
Åsele	Västerbotten	Norrland	2	0	0	2
Skellefteå	Västerbotten	Norrland	3	0	0	3
Nederkalix	Norrbottnen	Norrland	1	0	0	1
Gammalsvenskby	(Ukraine)		1	3	0	4
<i>Standard Swedish</i>			4	0	0	0

4.3. S-articles in the Swedish dialects

Proceeding now to the S-article, we saw earlier that the use is more restricted in Danish than in Norwegian and Swedish, in that in Danish the S-article is in complementary distribution not only with P-articles but in general with preposed determiners. As for the cases which distinguish Danish from the two other languages, there is essentially no variation among the traditional Swedish dialects. All of them, including the historical Danish provinces, consistently use double definite articles and S-articles also with demonstratives. It should be pointed out here that the same pattern is found in the Danish island of Bornholm, where the dialect is close to that of the neighbouring Swedish province of Skåne, suggesting that double determination in the southern Swedish area may have arisen already under Danish rule (before 1658). The interesting variation within Sweden with respect to the S-article is found elsewhere, in an area which I shall for convenience call “the peripheral Swedish dialect area”, and which includes the part of Dalarna where Dalecarlian is spoken, most of Norrland, Gotland, Finland, Estonia, and even the northern part of Värmland. In the peripheral area, the definite article has undergone an expansion process, as a result of which it is used in a number of contexts where it does not normally occur in the standard Scandinavian languages and in fact partly contradicts common ideas of what a well-behaved definite article should look like. The picture is somewhat complicated, since there are a fairly

large number of different uses, each of which seems to have a geographical distribution of its own. It is therefore impossible to give the whole story in this paper, and the essential point here is anyway that there is a general tendency towards an expansion of the S-article in the area mentioned, which is by and large not matched anywhere else in Scandinavia. I shall thus review the different extended uses rather briefly.

4.3.1. *Generic uses*

It is sufficient to compare English and French to see the use of definite articles in generic noun phrases may vary quite widely between languages.

Thus in English, definite articles are not in general used with generic NPs (excluding the type *The tiger is a wild animal*) whereas in French, where definite articles are virtually always used in those contexts. The standard Scandinavian languages are somewhere in between English and French in this respect. Thus, Swedish follows French in using a definite NP in *Livet är kort* ‘Life is short’ (cf. French *La vie est brève*) but like English prefers a bare noun in *Guld är dyrt* ‘Gold is expensive’. However, many dialects in the peripheral area seem to behave more like French in this respect. We thus find examples such as the following:

- (27) Älvdalen, Dalecarlia
Guldiö *ir* *dyrt.*
 gold: DEF be: PRS.SG expensive: N
 ‘**Gold** is expensive’

Related are examples like (28), where the name of an illness is used with an S-article:

- (28) Östmark, Värmland
Han *ä* *ill* *kommen* *ta* ***jekta.***
 he be: PRS badly come: PP of gout: DEF
 ‘He suffers badly from **gout**’
 (cf. Standard Swedish *Han lider illa av gikt*)

As generic one might also count uses of nouns with S-articles as “citation forms”, which seems to be quite common in many parts of the peripheral

area. Thus, speakers who are asked to write down word lists often quote nouns in the definite form.

4.3.2. Non-delimited (“partitive”) uses

The uses that I shall here call “non-delimited” have often been called “partitive” in the literature, which is good in the sense that they occur when Finnish has the partitive case and French uses the “partitive article”. As pointed out in Koptjevskaja-Tamm (2001: 525), however, the term “partitive” had better be reserved for constructions that express a part-whole relationship. For constructions that derive historically from partitive constructions but are rather used to express a non-specified quantity of something, she uses the term “pseudo-partitive”. This term is less suitable for patterns that have basically nothing to do with partitive constructions in the proper sense, and I therefore propose to call these uses “non-delimited”. The heads of non-delimited NPs are either mass nouns or plural count nouns. In English, they would typically be “bare NPs”, e.g. *beer* in *I am drinking beer*.

While generic NPs are grammatically definite in many languages, it is more astonishing to find that non-delimited NPs occur with definite articles. Yet this is the normal pattern in East Dalecarlian, the dialects in the northern part of Norrland (Jämtland, Ångermanland, Västerbotten, and Norrbotten) and Österbotten in Finland. In the Cat Corpus we thus find:

- (29) Våmhus (Dalarna)
Ja, bara i a faið in wi:ðn,
 yes only I have: PRES get: SUP in firewood: DEF
so ska i werm miö: tsī a na.
 so shall I warm: INF milk: DEF for PRO.3SG.F.DAT
 ‘[Granny says:] As soon as I have put in the firewood, I’ll warm
some milk for it (the cat)’ (Cat Corpus)

- (30) Skellefteå (Västerbotten)
Hä gick skaplit att klaar sä,
 it go: PST okay to survive

meda 'ä fanns rått-än å mus-än...
 as long as it exist: PST rat-DEF.PL and mouse-DEF.PL
 ‘[The cat thinks:] It was kind of OK [to live in the forest], as long
 as there were **rats and mice...**’ (Cat Corpus)

Notice that in (30), the definite forms are used in the “presentation construction” with the dummy subject *'ä*, which is normally restricted to indefinite NPs.

Slightly different are “weather expressions” such as the following:

- (31) Luleå (Norrbotten)
He er duvinver-e idä.
 it is muggy_weather-DEF today
 ‘The weather is muggy [lit. It’s **muggy weather**] today’

It is plausible that non-delimited uses of definites historically are extensions of generic uses. One argument in favour of this is that in the peripheral area, the generic uses have a wider geographical distribution which seems to properly include that of the non-delimited uses. Typological parallels are not so common but do exist, e.g. in Magrebian Arabic.

- (32) Moroccan Arabic
Kāin əl-ħobz.
 there-is DEF-bread
 ‘There is **bread**’ (Caubet 1983: 235)

4.3.3. Quantified NPs

In a more restricted area in Upper Norrland it is also possible to use S-articles after quantifiers such as ‘much’ and numerals, for instance:

- (33) Sorsele (Västerbotten)
Han ha tre brödren.
 he have: PRS three brother: DEF.PL
 ‘He has **three brothers**’

- (34) Nederkalix (Norrbotten)
Wå söynda skå dö vä så mytji sopan?
 what sin: DEF shall you with so much milk: DEF
 'What the hell are you going to do with so **much milk**?' (Cat Corpus)

4.3.4. Low referentiality uses

In the peripheral area, there are also some unexpected uses of S-articles with singular count nouns. In English, these would in general have an indefinite article, but in the standard Scandinavian languages they are typically realized as “bare nouns”. I call this “low referentiality uses”, since a common trait for them is that the identity of the referent is not highlighted. Typical examples would be *Vi har hund* ‘We have a dog’ and *Vi har telefon* ‘We have a telephone’, which could also be paraphrased as ‘We are dog-owners’ and ‘We are telephone-owners’, respectively. In the latter case, a definite article may be used in French: *Nous avons le téléphone*. Standard grammars of French call this a “generic article”, but that is a somewhat liberal use of that term. With respect to the peripheral Swedish dialect area, low referentiality uses are not well documented in the literature, and it is not so easy to establish the precise geographical distribution. Here is a clear example from the Cat Corpus:

- (35) Nederkalix (Norrbotten)
Jä skå tåla åom för dä, måmme,
 I shall speak: INF about for you: DAT mother
åt jä ållti veillt hå i kjaatt
 that I always want: SUP have a cat
män hä gja jo ät håå kjatta
 but it go: PRES as_you_know not have: INF cat: DEF
når man båo ini i höreshöus.
 when one live: PRS in a rent-house
 ‘I want to tell you, Mother, that I have always wanted to have a cat – but it isn’t possible to have a cat (lit. **the cat**) when you live in an apartment house’ (Cat Corpus)

Certain types of low referentiality uses, notably in prepositional phrases, are found also in southern Finland and Estonia, where non-delimited uses do not in general show up.

4.3.5. *Possessive NPs*

Within the peripheral area, there is an extreme diversity in the ways possessive noun phrases are expressed. Even in one and the same dialect, several different possessive constructions may compete with each other. What interests us here is the use of S-articles in constructions that are otherwise identical to the standard Scandinavian genitive construction, that is, where the possessor NP is preposed to the head noun and is marked by the genitive (or possessive) suffix *-s*. In large parts of the northern area, the *s*-genitive is rare or lacking, and such examples as can be found appear to be influenced by the standard language (Delsing 1996). In Finland, Estonia and Gotland the *s*-genitive is quite common, but normally the head noun is marked by the S-article rather than zero-marked, as in the standard languages. Also in dialects along the coast of Norrland the same construction is found – it appears to be strongest in the south (the province of Hälsingland). The distribution of the S-article in possessives is thus rather different from that of the other extended uses of S-articles discussed here. It is possible that it is a relatively recent phenomenon in the north, where other possessive constructions are strong (e.g. with dative-marked possessors).

We thus find a plethora of extended uses of S-articles in what I have called the peripheral area. In spite of the somewhat confusing isogloss patterns, the phenomena in question show uniformity between dialects that are geographically very distant from each other, such as Älvdalen in northern Dalarna, Överkalix in Norrbotten and Runö in the Gulf of Riga. Moreover, these dialects are also the ones that most consistently exhibit conservative traits inherited from Old Scandinavian. Historically, the areas in question largely coincide with the agricultural expansion in the early Middle Ages (which in Sweden means the period 1050–1350). In my opinion, the roots of the developments that led to the expansion of the S-articles are to be found in central Sweden, in the provinces of Uppland and Södermanland. The reason why there are hardly any traces of the phenomena in those parts would then be that the urbanization process that started in the middle of the 13th century led to what was in effect the

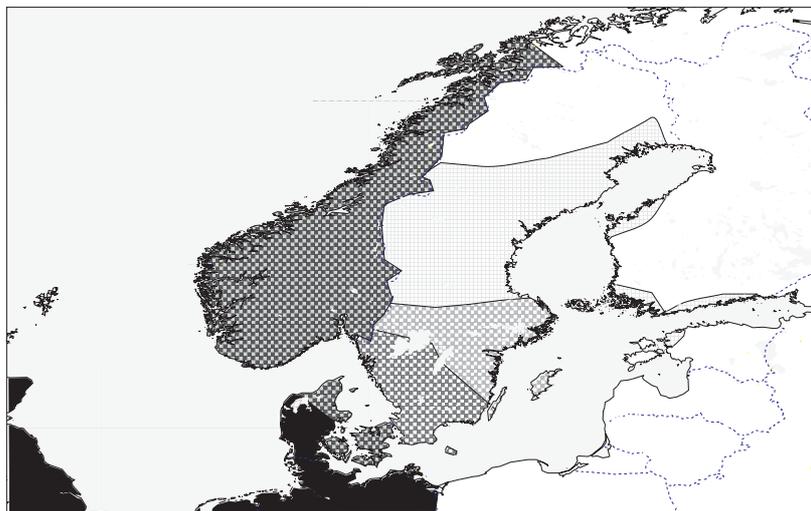
replacement of the dialects spoken there by varieties of a more southern character. From the point of view of this paper, this is however a side-issue. What is important here is that we find in the north-eastern part of the Scandinavian dialect continuum area a further grammaticalization of S-articles beyond what we think of as normal for definite articles. I say “what we think is normal”, since it is well attested that definite articles may expand their domains of use to a point where they are obligatory noun affixes, as in the Bantu languages, where definite articles are assumed to be the source of the class prefixes (Greenberg 1978). In north-eastern Scandinavia, we find an intermediate stage of this development of a kind that has not been well documented cross-linguistically.

5. Conclusions

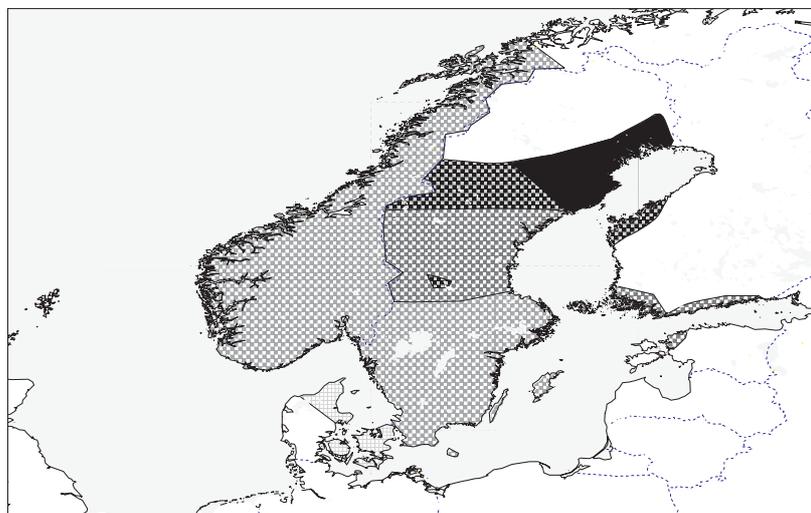
With regard to P-article use, Standard Swedish turns out to be intermediate between the traditional dialects in the south, where the P-article is strong, and the dialects in the north and east of the Baltic, where the use is weak. This obviously is consistent with the general picture. The Continental Scandinavian dialect continuum area may be said to consist of five different sub-areas, with respect to the strength of the P-article, ordered in decreasing strength from the south-west towards the north-east as follows: (i) SW Jutland (where the P-article is in general use as a definite article), (ii) the rest of Denmark; (iii) Norway and southern and western Sweden; (iv) central Sweden; (v) northern Sweden (see Map 3).

It is possible to make a similar ordering of sub-areas with respect to the strength of the S-article, but it turns out to be more or less the inverse of what we saw for the P-article (see Map 4). The widest use of the S-article is found in the north-east, in the “peripheral Swedish dialect area”. Central and southern Sweden and Norway constitute an intermediate zone, while Denmark displays the strongest restrictions on the use of the S-article. In fact, as we approach the limit of the S-article area, these restrictions become even stronger. On the very borderline between the S-article and the more general, “Western European” pattern of P-article use, not only adjective-noun combinations but also single nouns are used without any article at all in what was called “name-like uses” above. (Such uses are basically restricted to kin terms in the standard languages.) Parallels can be found also in Icelandic and Faroese, indicating that the grammaticalization of S-articles there has not reached the point that it has in most of the

continental Scandinavian area. However, in these geographically peripheral languages, situated opposite of the peripheral dialects in the north-east, the P-article is also less well developed – in modern spoken Icelandic it seems wholly absent.



Map 3. The geographical distribution of P-articles



Map 4. Geographical distribution of S-articles

The two main definite article types in Scandinavia – called here “P-articles” and “S-articles” – thus represent two separate grammaticalization processes represented in overlapping geographical areas with different “centres of gravity”. The S-article (historically oldest in the area) has reached its fullest development in north-eastern Scandinavia. The P-article (a newcomer from the south) has been frozen at a relatively early stage and has left some areas virtually untouched. In this perspective, it is natural to see “double determination” as one possible outcome of the competition between two different grammaticalizing definite articles. Another solution is chosen in Standard Danish, where the two articles are in a perfect complementary distribution. But we can note that Swedish is more prone than Norwegian to choose a one-article solution, but normally in favour of the S-article.

To see all this in the right light, we need to look not only at the standard Scandinavian languages, but also take into account the astonishing diversity found in the traditional dialects and also try to understand the diachronic developments behind the synchronic patterns, using information about earlier stages of the languages and the findings of typological and diachronic linguistics, in particular what concerns possible grammaticalization paths.

Acknowledgment

The work reported here has been supported by a grant from the Bank of Sweden Tercentenary Foundation.

Notes

1. For another typologically oriented discussion of double determination in the Scandinavian languages, but from a slightly different point of view, see Börjars (1994). The most systematic treatment of dialectal variation in the syntax of the noun phrase in Scandinavian is found in Delsing (1996). Plank (2003) is a survey of the phenomenon of “double articulation”, defined as violations of the expectation that at most one definiteness or indefiniteness marker will be used in each NP.

As for the general diachronic development of definite articles, the typical source is distal demonstratives (‘that’). It appears that the first uses to appear are anaphoric, followed later by “situational uses” such as *I wanted to open the door but could not find **the keys*** (where the keys have not been mentioned

before but are identifiable in the context). Later on, definite articles may come to be used with generic noun phrases and proper names, and in some cases, may extend further to uses which would not normally be thought of as definite and finally end up as a general marker on nouns. As is common in grammaticalization processes, an article may at some stage of its development lose its status as a word and become an affix. See further Lyons (1999: 322–340).

2. I have not tried to separate out the two standard language varieties Bokmål and Nynorsk here, but there do not seem to be any essential differences.
3. The representativity of Internet searches may of course be questioned. I have used them as a heuristic instrument but have in general refrained from giving absolute numbers since these would demand a much more careful checking of individual examples. For the crucial cases where the languages differ, the differences in numbers are large enough to be seen as reliable.
4. An Internet search reveals that the differences between the two varieties of Norwegian are not so clear-cut as normative grammars would have it, in this case. Thus, both *det Hvite huset* and *det Kvite hus* do occur. It appears that the pattern P+W+S– has its source in Danish-inspired written language and is not productive in spoken language.
5. The speakers of this dialect are descendants of Swedish-speaking peasants from Estonia who were re-settled to Ukraine in the 18th century. In 1929, most of them emigrated to Sweden and Canada, but a few still remain. The translation into Gammalsvenska was made by emigrants in Sweden.
6. The dialects in the provinces of Värmland and Östergötland are usually regarded as “Göta” dialects but exhibit a rather heterogeneous behaviour. In this investigation, they do not line up with the other dialects from the Southern and Western area and I have labeled them “transitional” in Table 1.
7. My account here glosses over various problems – these issues need to be worked out further.

References

- Börjars, Kersti
 1994 Swedish double determination in a European typological perspective. *Nordic Journal of Linguistics* 17: 219–252.
- Caubet, Dominique
 1983 Quantification, négation, interrogation: les emplois de la particule “šī” en arabe marocain. *Arabica* 30: 227–245.

- Dahl, Sven
 1989 Studier i äldre skånska odlingssystem [Studies in older Scanian cultivation systems]: *Meddelanden från Kulturgeografiska institutionen vid Stockholms universitet*, B 69. Stockholm: Kulturgeografiska institutionen, Stockholms universitet.
- Delsing, Lars-Olof
 1996 Nominalfrassyntax i skandinaviska dialekter [Noun phrase syntax in Scandinavian dialects]. *Nordica Bergensia* 6: 24–74.
- Greenberg, Joseph H.
 1978 How does a language acquire gender markers? In: Joseph H. Greenberg (ed.), *Universals of Human Language*, 48–81. Stanford, CA: Stanford University Press.
- Koptjevskaja-Tamm, Maria
 2001 “A piece of the cake” and “a cup of tea”: Partitive and pseudo-partitive nominal constructions in the Circum-Baltic languages. In: Östen Dahl and Maria Koptjevskaja-Tamm (eds.), *The Circum-Baltic Languages: Typology and Contact*. Volume 2: *Grammar and Typology*, 523–568. Amsterdam/Philadelphia: John Benjamins.
- Lyons, Christopher
 1999 *Definiteness*. Cambridge: Cambridge University Press.
- Plank, Frans
 2003 Double articulation. In: Frans Plank (ed.), *Noun Phrase Structure in the Languages of Europe*, 337–396. Berlin/New York: Mouton de Gruyter.
- Teleman, Ulf, Staffan Hellberg, Erik Andersson, Lisa Christensen and Svenska Akademien
 1999a *Svenska akademiens grammatik. 2. Ord* [Swedish Academy Grammar. 2. Words]. Stockholm: Norstedts.
 1999b *Svenska akademiens grammatik. 3. Fraser* [Swedish Academy Grammar. 3. Phrases]. Stockholm: Norstedts.