Grammaticalization
of (in)definiteness in Swedish
Dominika Skrzypek

Grammaticalization of (in)definiteness in Swedish
Modern Swedish is an article language which in the course of its history has developed both definite and indefinite articles. The present study focuses on the grammaticalization of the postposed definite -IN, etymologically the demonstrative pronoun hinn 'yon' and the indefinite EN, whose source is the numeral 'one'. Both grammaticalizations originate with textual uses of the forms: the demonstrative to mark anaphora and the numeral to mark new, persistent referents. The processes are interdependent: the indefinite article does not grammaticalize before the grammaticalization of the definite has been triggered and its development restricts the spread of the definite article to specific uses.

Since both grammaticalizations have the same sources in unrelated article languages, this detailed study of the developments in Swedish has a bearing on the studies of other languages as well.

Key words: grammaticalization, definite article, indefinite article, grammaticalization chai

Dominika Skrzypek, Department of Scandinavian Studies at the Adam Mickiewicz University, al. Niepodległości 4, 61-874 Poznań, Poland

Recenzent: dr hab. Janusz Stopyra

Praca powstała w ramach projektu badawczego habilitacyjnego N N104 461740
Wydano na podstawie maszynopisu gwarantowanego

© Dominika Skrzypek 2012

This edition © Uniwersytet im. Adama Mickiewicza w Poznaniu,
Wydawnictwo Naukowe UAM, Poznań 2012

Projekt okładki: Ewa Wąsowska

Fotografia na okładce: najstarszy zachowany szwedzki manuskrypt zawierający Äldre Västgötalagen (Starsze prawo Gautolandii Zachodniej) z XIII wieku, B 59 w zbiorach Biblioteki Królewskiej w Sztokholmie

Redaktor techniczny: Elżbieta Rygielska
Łamanie komputerowe: Anna Marcinkaniec

ISSN 0208-7235
## Contents

Abbreviations .................................................................................................................... 8  
Acknowledgements ............................................................................................................ 9  

### 1. Introduction ............................................................................................................. 11  
1.1. Statement of intent ................................................................................................. 11  
1.2. Periodisation of the Swedish language .................................................................... 13  
1.3. The sources .............................................................................................................. 15  
1.4. The method .............................................................................................................. 21  
1.5. (In)definiteness in Modern Swedish ........................................................................ 23  
  1.5.1. NP types in MS ..................................................................................................... 23  
  1.5.2. The functions of the definite article in MS ......................................................... 27  
  1.5.3. The functions of the indefinite article in MS ...................................................... 30  
  1.5.4. Use of bare nouns in MS .................................................................................... 32  
1.6. Organisation of the book ......................................................................................... 34  

### 2. Grammaticalization and (in)definiteness .................................................................. 35  
2.1. Introductory remarks ............................................................................................... 35  
2.2. What is grammaticalization? ................................................................................... 35  
2.3. Major tenets of grammaticalization ........................................................................ 38  
  2.3.1. Changes in meaning ............................................................................................ 38  
  2.3.2. Changes in form ................................................................................................ 40  
2.4. The role of context in grammaticalization .............................................................. 43  
2.5. Grammaticalization chains ..................................................................................... 44  
  2.5.1. From a demonstrative to a definite article ......................................................... 45  
  2.5.2. From numeral ‘one’ to an indefinite article ....................................................... 51  
  2.5.3. Grammaticalization of (in)definiteness–a large chain? ....................................... 54  
2.6. Grammaticalization and language change .............................................................. 55  
2.7. Conclusions ............................................................................................................. 55  

### 3. Sources of grammaticalization ................................................................................. 57  
3.1. Introduction ............................................................................................................. 57  
3.2. Etymologies of the articles ..................................................................................... 57  
  3.2.1. The weak adjectival inflection .......................................................................... 58  
  3.2.2. Demonstratives .................................................................................................. 62  
  3.2.3. Numeral ‘one’ .................................................................................................... 65
4. Grammaticalization of the definite article........................................................................... 78
  4.1. Introduction .................................................................................................................. 78
  4.2. General results ............................................................................................................. 78
  4.3. Åldre Västgötalagen (AVL) .......................................................................................... 81
    4.3.1. Textual uses—direct anaphora in AVL ................................................................. 82
    4.3.2. Textual uses—indirect anaphora in AVL ............................................................... 85
    4.3.3. Non-textual functions in AVL ............................................................................... 88
    4.3.4. Distribution of -IN forms in AVL ......................................................................... 90
    4.3.5. Excursus: on the form ‘arwumn’ in AVL and YVL ............................................... 91
  4.4. The development of textual functions ......................................................................... 92
    4.4.1. Direct anaphora ..................................................................................................... 92
    4.4.2. Indirect anaphora ................................................................................................. 107
    4.4.3. Excursus: inalienable possession ........................................................................ 116
  4.5. The development of non-textual functions ................................................................ 121
    4.5.1. Uniques ................................................................................................................ 121
    4.5.2. Generics ............................................................................................................... 133
    4.5.3. From textual to non-textual uses of -IN .............................................................. 135
  4.6. Anomalous uses of -IN .............................................................................................. 137
    4.6.1. Bondi in legal prose ............................................................................................. 137
    4.6.2. Co-occurrence of -IN with strong forms of adjectives ......................................... 145
    4.6.3. Co-textual uses of -IN in religious prose ............................................................. 146
    4.6.4. Co-occurrence with numeral EN ........................................................................ 151
  4.7. Discussion of the distribution of -IN .......................................................................... 152
  4.8. Conclusions ................................................................................................................ 154

5. Grammaticalization of the indefinite article ..................................................................... 155
  5.1. Introduction ................................................................................................................ 155
  5.2. General results ............................................................................................................ 156
  5.3. Numeral EN .............................................................................................................. 158
  5.4. Some remarks on specific and nonspecific uses of EN .............................................. 160
  5.5. Presentative EN ....................................................................................................... 161
  5.6. Specific EN ............................................................................................................... 170
  5.7. Non-specific EN ....................................................................................................... 174
    5.7.1. Comparison ........................................................................................................ 175
    5.7.2. Negation ............................................................................................................. 179
    5.7.3. Conditionals ....................................................................................................... 180
    5.7.4. Verbs of volition ............................................................................................... 182
    5.7.5. Predicatives ...................................................................................................... 182
  5.8. EN vs NGN .............................................................................................................. 184
  5.9. Anomalous uses of EN: EN defNP ........................................................................... 190
  5.10. Conclusions ............................................................................................................. 193
6. Conclusions .................................................................................................................... 194
  6.1. Introduction ............................................................................................................. 194
  6.2. Overview of the grammaticalization of the definite and indefinite article in Old Swedish .......................................................... 195
    6.2.1. Summary of the developments .................................................................... 195
    6.2.2. Textual origins of the articles ...................................................................... 196
    6.2.3. Stages of the grammaticalization of the articles in OS ......................... 197
    6.2.4. Note on the co-textual definiteness .............................................................. 198
  6.3. Discussion of the models of grammaticalization ................................................... 198
  6.4. Interplay between the grammaticalizations ............................................................ 200
  6.5. Interplay with other grammatical changes ............................................................ 202
  6.6. The long-term perspective ...................................................................................... 203
  6.7. Concluding remarks ............................................................................................... 203

Appendix: Notes on NP types in OS ................................................................................. 205
Sources ............................................................................................................................... 209
References .......................................................................................................................... 211
List of tables and list of diagrams .................................................................................... 221
Index .................................................................................................................................. 223
Gramatykalizacja (nie)określoności w języku szwedzkim (streszczenie) ....... 226
Abbreviations

Languages

AN – Ancient Nordic
EN – East Nordic
MS – Modern Swedish
IE – Indo-European
OI – Old Icelandic
ON – Old Nordic
OS – Old Swedish
PGmc – Proto-Germanic
PIE – Proto-Indo-European
WN – West Nordic

Grammatical terms

The suffix -in, with sources in the demonstrative (h)inn is referred to as -IN. En, originally numeral ‘one’, the source of the indefinite article, is referred to as EN. NGN stands for the pronoun någon ‘some/any’.
Apart from two very long examples, all Old Swedish and Old Icelandic examples are glossed, but the grammatical information is kept to the minimum.

ACC – accusative
defNP – definite noun phrase
DAT – dative
indefNP – indefinite noun phrase
DEF – definite
PART – particip
DET – determiner
PASS – passive
DO – direct object
PL – plural
FUT – future
PP – prepositional phrase
GEN – genitive
PRET – preteritum
INDEF – indefinite
REFL – reflexive
IO – indirect object
SG – singular
N – noun
ST – strong adjectival inflection
NOM – nominative
W – weak adjectival inflection
NP – noun phrase

Abbreviations of source text names can be found under Sources at the end of the book and in Chapter 1.3
Acknowledgements

I am greatly indebted to a number of people without whom this book could never have been written.

Professor Witold Maciejewski has long supported me in my research, first and foremost as the supervisor of my MA and PhD thesis, but even after, making me realize how much there is to learn and find out about languages long after this formal role was over. He has helped me in more practical matters as well, assisting in many applications for research funds. His encouragement and friendship have meant a great deal to me in the course of the work and he is the main source of inspiration for many of the ideas presented in this book.

I have had the privilege of discussing my work with two people who are the foremost experts on the history of the Swedish language and historical linguistics: Professor Cecilia Falk and Professor Lars-Olof Delsing. Both have kindly extended invitations for me to complete parts of my research in Stockholm and Lund respectively. Apart from the great scholarly opportunities this gave me, I could also enjoy the milieu of the universities with the wide selection of seminars and courses offered there. These two long stays in Sweden would not have been possible without the financial support of the Swedish Institute, who have kindly assisted me in this.

The results given here have been presented and discussed on conferences Svenska språkets historia 10 and 12 in Gothenburg and in Stockholm, as well as on a number of seminars in Stockholm, Lund, Poznań, Warsaw and Greifswald. I would like to thank all the participants for the constructive critique and ideas for improvement. In particular, I would wish to thank professor Valéria Molnár of Lund university for the inspiring discussions and seminars and professor Folke Josefsson for the discussion of my research at Svenska språkets historia 10 in Gothenburg.

I would also like to thank Professor Halldór Ármann Sigurðsson who has helped me with the Old Icelandic examples. His courses have also sparked the interest for the Icelandic language, making it somewhat more prominent in the study than originally intended.
I thank Siddharth Narayan for the proofreading of the text and ruthless eradication of all ‘interestingly enough’ which, interestingly enough, seemed to thrive in almost every paragraph.

More informally, my dear friends at the Institute of Scandinavian studies in Poznań are to be appreciated for the animated discussions of both historical linguistics and the difficult fate of a humanist in the modern world. Joanna, Paulina, Dorota, Sylwia, Aldona, Mikołaj and Piotr have always helped me through hectic times, unfailingly giving the moral support needed.

Finally, I would like to thank my parents, who have shown me the joys of discovering the world and making the discovery a profession!

Poznań, May 2012

Dominika Skrzypek

The research presented in this book was financed by two grants from the Swedish Institute and by research grant from the Polish Ministry of Science and Higher Education, grant no N N104 461740.
The grammaticalization of definite articles is a relatively common phenomenon within languages, and though not entirely uniform, also shows cross-linguistic similarities among unrelated languages. Grammaticalization of the indefinite article, on the other hand, is much rarer, though it also shows similar cross-linguistic patterns.

Almost 2/3rd of all languages in the WALS inventory have definite articles (that are distinct from demonstratives), but only ca 1/3 have indefinite articles (including, both, articles different from and identical to the numeral ‘one’). Also, while it is likely that a language will have the definite article if it has an indefinite article, the reverse does not hold, see table 1.

Table 1. Distribution of definite and indefinite articles in the world’s languages

<table>
<thead>
<tr>
<th>Articles in the world’s languages</th>
<th>No of languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both definite and indefinite article</td>
<td>194</td>
</tr>
<tr>
<td>No definite or indefinite article</td>
<td>198</td>
</tr>
<tr>
<td>No indefinite, but definite article</td>
<td>98</td>
</tr>
<tr>
<td>No definite, but indefinite article</td>
<td>45</td>
</tr>
</tbody>
</table>

Source: *WALS*

It also seems that the indefinite articles develop after the grammaticalization of the definite ones has at least been initiated; such is the case in the Romance and Germanic languages, but cross-linguistic data show that an article language typically possesses either the definite article, or both definite and indefinite articles, but seldom only the indefinite. Studies of the grammaticalization of either definite or indefinite article so far also show that articles are a relative novelty in
languages, and rather than having to be reconstructed many of these developments can be studied in the actual texts.

The aim of the present study is to map the developments of the definite and indefinite articles in Swedish, a language representing the Scandinavian languages, the Northern branch of the Germanic languages, with particular attention paid to the interdependencies between these developments. Definiteness marked overtly by definite and indefinite articles is not part of the common Indo-European (IE) heritage, nor is it found in Proto-Germanic (PGmc). And although all modern Germanic languages have articles, their rise belongs to their separate histories. With respect to definiteness Swedish is particularly interesting since it has both a bound definite article and a free-form one, as well as the so-called weak adjectival declension (see 1.5). The bound definite article is a suffix of a different origin to the (preposed) definite articles in e.g. English and German (see Chapter 3). Swedish also has the indefinite article whose development can be studied in its entirety in the Old Swedish (OS) sources.

The grammaticalization of the definite and indefinite articles in Swedish will be presented on the basis of a detailed study of a sample of OS texts. The empirical goal is to illustrate the development of the original forms: demonstratives *(h)iinn ‘that’, *sá ‘this’ and numeral *en ‘one’ into articles. There will be focus on the spread of the forms from one function to the next. The study also has a two-fold theoretical goal: to validate and complete the models of grammaticalization of definite and indefinite articles as well as to connect the two developments to one another.

Modern Swedish (MS), like its close relatives Danish, Norwegian, Icelandic and Faroese, is different from other Germanic languages in that its definite article is a suffix and not a preposed lexeme; thus the English defNP *the man* or the German *der Mann* corresponds to the Swedish *mann-en*. The suffix is a descendant of the Old Nordic (ON) demonstrative *hinn* ‘that’/‘yon’. In defNPs that contain an adjective apart from the noun, the adjective takes a weak form and is preceded by the definite determiner *den*, a continuant of the ON *sá/pen* ‘this’. The result is therefore e.g. *den glad-a mann-en* ‘the happy man’. The Swedish indefinite article is, like in other Germanic languages that have developed it, a descendant of the numeral ‘one’, in ON *einn.*

---

1 Some authors use the term suffix (e.g. Faarlund 2007, Börjars and Harris 2008) and some clitic (Syrett 2002). For a discussion of the status of -IN as either clitic or affix in Old Nordic, which is valid even for OS, see Faarlund 2007, Abraham and Leiss 2007 and a critique of both in Börjars and Harris 2008.

2 For periodisation and terminology, see 1.2.

3 The original nominative masculine *sá* is in OS substituted by the accusative *pen*.

4 Note that there is a difference between Swedish and its close relative, Danish, which in such defNPs does not use the suffix, thus Danish counterpart to ‘the happy man’ is *den glade mand*.

5 Among those that have not is one Scandinavian language, namely Icelandic.
This study will address the following questions:

1. Functionally and pragmatically, how do (h)inn and en proceed from their original functions as a demonstrative and a numeral respectively to these of articles? Which article functions appear first, and which lag behind? Are there any constraints on these developments?

2. Can the clines of grammaticalization proposed in literature be observed in the Swedish material? Can the developments be outlined in more detail?

3. Is there any interdependence between the grammaticalization of the definite article and the grammaticalization of the indefinite? What is their relative chronology?

A frequently asked question is why. Why does a change occur? What is its instigation? For all the ‘whys’, the ‘how’ often remains a secret. When it comes to the development of definite and indefinite articles in the Scandinavian languages, the why and when have dwarfed other considerations (see Chapter 3.3 for an outline of the earlier studies). Therefore a solid description of the process is lacking. The present study will hopefully fill this void.

### 1.2. Periodisation of the Swedish language

Swedish belongs to the Scandinavian languages, the northern branch of the Germanic languages. The separation of the northern branch from the rest of the Germanic languages most likely started in the early centuries AD (Bandle et. al. 2002), but the language first took form around 500 AD. Around 700-800 internal differences increased in significance (and can be seen in the runic inscriptions from that period), gradually splitting the language into the western and eastern varieties.

Diagram 1. Periodisation of the North Germanic languages before 800 AD

<table>
<thead>
<tr>
<th>Period</th>
<th>Language</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>200-450/500</td>
<td>Ancient Nordic</td>
<td>AN</td>
</tr>
<tr>
<td>450/500-700/800</td>
<td>Ancient/Old Nordic</td>
<td>AN/ON</td>
</tr>
<tr>
<td>700/800-1100</td>
<td>Old Nordic</td>
<td>ON</td>
</tr>
<tr>
<td>(ca 800 split into)</td>
<td>– West Nordic</td>
<td>WN</td>
</tr>
<tr>
<td></td>
<td>– East Nordic</td>
<td>EN</td>
</tr>
</tbody>
</table>

*(based on Bandle et al. 2002)*
Until ca 1200 the writing system used in Scandinavia was the fuþark, the runic alphabet, related to but distinct from the Mediterranean alphabets. There are few inscriptions from the earliest centuries, but from ca 800 AD runic sources become more numerous. Since they include a number of inscriptions carved in stone and not on small, transportable objects like pieces of wood or metal, many can be placed in Sweden and treated as the earliest sources of the language spoken on the Swedish territory. It has therefore become customary to refer to the time 800-1200 AD as Runic Swedish.

The traditional periodisation of the history of the Swedish language is based on factors external to language, mainly the publications of important texts. Thus the Old Swedish period (fornsvenska) begins with Äldre Västgötalagen (AVL, legal text dated at 1225 and the first text written in the Latin alphabet) and ends 1526 with the first complete translation of the New Testament into Swedish. The New Swedish period (nysvenska) which then begins is divided into the older nysvenska, ending 1732 when the influential Den svenske Argus was published, and younger nysvenska. Naturally all these texts are of great consequence for the development of the language. However such exact dates cannot possibly be considered dates of stages in linguistic development.

The period of interest for this study and the period of greatest change in the documented history of the Swedish language is Old Swedish. It is further subdivided into older (or classical) and younger, with an approximate border ca 1375 (e.g. Wessén 1941). 1375 is a sensible proposition from the purely linguistic point of view. The texts written before are in a clearly conservative language whereas texts from the second part of the Old Swedish period show great reductions in the original morphological system. These include, among others, the loss of the case system, as well as the final stages of grammaticalization of the definite and the indefinite articles.

It has become customary in historical linguistic studies of Old Swedish to further subdivide the younger part of the period at 1450. Such subdivision is more true to the linguistic development (Hirvonen 1987, Håkansson 2008) and this practice will be followed here, resulting in three subperiods of OS studied separately.

<table>
<thead>
<tr>
<th>Period</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runic Swedish</td>
<td>800-1225</td>
</tr>
<tr>
<td>Old Swedish</td>
<td></td>
</tr>
<tr>
<td>Period I</td>
<td>1225-1375</td>
</tr>
<tr>
<td>Period II</td>
<td>1375-1450</td>
</tr>
<tr>
<td>Period III</td>
<td>1450-1526</td>
</tr>
</tbody>
</table>
1.3. The sources

The earliest documents of the Swedish language are runic inscriptions (ca 3000 have been discovered so far). Many of these are short and formulaic in character. The demonstrative *sá* and a compounded form *sjá (denna)* are richly represented, but only two doubtful instances of definite article and one of a possible indefinite can be found in the material (see Chapter 3.4). Runic inscriptions are here quoted with their ‘signatures’, e.g. U 136, where the letter is a symbol of the Swedish province where the inscription was found. All inscriptions can be found using the signature in the database provided by the Uppsala University (http://www.nordiska.uu.se/forskn/samnord.htm).

Sources from Period I are relatively few. These include mainly the legal codices for Swedish provinces of that time (excluding the then Danish Scania, Halland and Blekinge), a collection of religious legends known as Codex Bureanus (a translation of the famous Legenda Aurea), Pentateukparafrasen—a translation of the biblical Tora, a collection of versed tales translated from French romances (partly via Old Norwegian translations), the so-called Eufemiaavisorna, and a versed history of Sweden, Erikskrönikan.

In period II the first common law for the whole country was drafted (*Magnus Erikssons landslag*, 1350) and the translation of religious prose continued. A new addition in this period are the profane narratives, also translations, mainly from Latin, like *Sju vise mästare* ‘The seven sages’ (from ca 1400).

An overview of the texts studied can be found in Table 2. The second column gives the abbreviations which will be used for a given text here, the third the full titles and the fourth the date at which the text was written. The sources are presented from the oldest to the latest, it should however be borne in mind that most of the dates are speculative. Therefore column 5 gives the dates of the manuscripts on which the excerpted editions are based. In some instances (e.g. Pent) the discrepancies are large.

An important aspect is the length of the texts studied (shown in column 6). Since the differences between texts here are large (younger texts tend to be longer), the shorter texts were studied in extenso (including AVL, the oldest Swedish text of considerable length), while shorter passages were chosen from the longer ones. Additionally the frequency of the definite and indefinite forms rises, rendering more empirical material in smaller samples. The lengths of the chosen passages is given in column 7.
The oldest Swedish texts represent the following genres: legal, religious and profane prose and poetry. An attempt was made to give the first three genres full representation. Poetry was not included as the use of the definite forms there indicates substantial influence of Middle Low German (e.g. Haskå 1972). For similar reasons the religious narrative *Själenz Tröst* ‘The Comfort of the Soul’ was omitted.
For period I, the choice was made to include the oldest texts available apart from the runic inscriptions—mainly AVL, the most important document of the Swedish language at its earliest stages. To avoid analyzing dialectal rather than chronological variation (although by definition this is not entirely possible in historical studies), YVL was also studied as it represents the same region as AVL and is in many respects its continuation (Skrzypek 2009). To represent other genres, Bur was chosen as the oldest religious text. These were completed with the first translation of the Bible Pentateuk (Tora), Pentateukparafrasen.

From period II a choice of religious (HML, Jart, Bo) and profane prose (SVM, KM) was made, completed with JB, which contains legal documents within which elements of spoken language may be found (in records of court dealings).

Finally, for period III two types of texts were chosen: two religious, three profane. A detailed overview of each text is given below.

- **Period I (1225-1375)**

  **Äldre Västgötalagen (AVL)**

  AVL is the oldest text in Swedish written not with runes but with Latin characters. The oldest fragment which was found can be dated at 1225. The whole text is found in a later manuscript from ca 1280.

  AVL is a legal codex of the province of Västergötland. It is divided into smaller parts dealing with different types of criminal offences, principles of sale and inheritance of property and regulations of religious matters such as the payment of tithes.

  AVL is linguistically very archaic even in comparison with other legal texts from the same period. For this reason it is studied in its entirety (ca 15000 words) and the definite forms are presented in great detail in Chapter 4.

  **Yngre Västgötalagen (YVL)**

  YVL is a continuation of AVL from the same province, written down around 1280 and preserved in a manuscript dated at ca 1350. A number of passages are identical with AVL; it seems that the manuscripts of YVL and AVL are based upon a text that is now lost (Skrzypek 2009).

  The whole text is ca 26000 words long, the parts chosen for this study are ca 9200 words and include Kyrkobalken (KB, the church codex), Giptobalken (GB, the marriage codex) and þiuvabalken (þB, the larceny codex).
**Östgötalagen (OgL)**

OgL was written around 1280 and the earliest preserved copy can be dated at ca 1350. The text is considered to be one of the most modern among legal prose, linguistically (Ståhle 1967).

The text is ca 53500 words long. The part excerpted here is 10220 words long and includes *Kyrkobalken* (KB, the church codex), *Giptobalken* (GB, the marriage codex) and *Vadamal ok sara mal, hor, ran ok styld* (Vm, the felony codex).

**Dalalagen (DL)**

DL is also known as *Äldre Västmannalagen*; it is not clear whether it was valid in one province (Dalarna) or two (Dalarna and Västmannaland). It is preserved in one manuscript only, dated by Schlyter at 1318-1347 (Schlyter 1841), however one may assume an earlier date of its formulation, since its continuant, *Yngre Västmannalagen*, was written down ca 1350.

The whole text is ca 16260 words long, the parts chosen for this study are ca 10750 long and include *kristnu balkar* (KB, church codex), *kunungsezöre* (KE, royal codex), *manhelgis balkir* (MB, personal and property rights), *gipningen balkir* (GB, marriage codex), *um þiufnadh* (þB, on larceny), *þingbalkir* (TB, ting codex).

**Codex Bureanus (Bur)**

Bur is the oldest OS text apart from the legal codices. It contains hagiographical legends and is a free translation of *Legenda aurea* by Jacobus de Voragines. The translation was probably produced at some time between 1276 and 1307 (Jansson 1934:4), the manuscript is dated at 1350-70 (Ottelin 1900:4, Jansson 1934:94).

The text is ca 36200 words long and for the purpose of the present study a shorter part was chosen, pages 133-192, ca 9770 words long.

**Pentateukparafrasen (Pent)**

Pent is a translation of the five books of Moses (Tora), probably produced in Vadstena around 1330 (Klemming SFSS 9:1, 577, but Thorell 1959:xliiiif finds it probable that the translation was made by bishop Brynolf at the end of the 13th century). There are two manuscripts of Pent available today, the so-called A manuscript dated at 1430-50 (Klemming SFSS 9:1, 588) and B manuscript dated at 1526 (according to a note at the bottom of the text, Hesselman 1927:8-9). Both are probably copies of the same original, however it is the later copy that was more faithful, as the earlier one has modernised the language (Hirvonen 1987:52). Therefore the present study is based on Pent B from which pages 156-182 were excerpted (ca 10077 words).
• **Period II (1375-1450)**

*Järteckensboken (Jart) and Helga manna leverne (HML)*

Both texts are included in the same manuscript, the so-called Codex Oxenstiernianus, written down in Vadstena and dated by its scribe at 1385. While the date is probably correct for Jart, HML is said to be some years younger and was probably written closer to 1400 (Mattsson 1957:232, Hirvonen 1987:54-55). The most important indication of their relative age is the use of the letter <þ> in Jart whereas HML uses the more modern spelling <dh> or <th>.

Jart and HML are religious texts consisting of a number of relatively short tales; in Jart they relate miracles connected among others with the host (altar bread)—järtecken means ’signs’ or ’miracles’; HML is a collection of hagiographic tales—a translation of the Latin *Vitae Patrum*.

For the present study the following pages were chosen: from Jart 3-49 (ca 10620 words) and from HML 181-223 (ca 10000 words).

*Stockholms stads Jordebok (JB)*

JB is a collection of legal texts documenting court procedures in mediaeval Stockholm. The texts are short, some in Swedish and some in Latin, and were written more or less in the order in which they appeared before court. For each year some space was left for complementations and this was duly used, sometimes some years later (Hildebrand 1876:427). While this should not seriously influence the linguistic value of the material it is however important to bear in mind that even though a text may be dated to a particular year it may in fact have been written some years later.

JB is not available as an electronic file, therefore the choice of text had to be limited. For the purpose of this study a choice was made from texts from the earliest years (1420-1450) and includes the documents numbered 4, 6, 8, 11, 67, 75, 178, 180 in Hildebrand’s edition. They are quoted as e.g. JB 178. The total amount of words is ca 2400.

*Karl Magnus (KM)*

KM is an abridged and adapted translation into Swedish of two tales of Charlemagne, based upon the Norwegian *Karlamagnús Saga*, a prose translation of a number of French poems about the king (a detailed filological analysis is given in Kornhall 1959). It has been preserved in four copies, all from the 15th century. The one that is the basis of the present study is found in a manuscript dated at 1430-50 (Geete 1903:127), some years after the translation to Old Swedish was first made (Ståhle 1967:109 gives ca 1400 as the most probable date). The tales
describe Charlemagne’s journey to Jerusalem and Constantinople and the Battle of Roncevaux. The text is about 10940 words long and was studied in its entirety.

\textit{Sju vise mästare (SVM)}

SVM is a narrative text with a Chinese box structure, where the main character is tempted by his wife to give orders to execute his son who is protected by the seven wise men responsible for his education. Both sides tell the king short stories with a moral that is to convince him of either sparing or slaying the boy. This means that there are two levels in this story—partly short novellas and partly the superordinate tale.

SVM is a translation from Latin (manuscripts A and B, Ståhle 1967:116) and German (manuscript C, Blomqvist 1941:263-264), available in three different manuscripts, all of them at least partly incomplete. The most comprehensive one, manuscript A, is dated at 1430-50 (Geete 1903:128), though the original translation is some 50 years older (Hirvonen 1987:58). The text is about 16300 words long and was studied in its entirety.

\textit{Bonaventuras betraktelser över Kristi liv (Bo)}

Bo is the OS translation of \textit{Meditationes vitae Christi}, written by cardinal Bonaventura (1221-1274). The text contains meditations on Christ’s childhood, with stories loosely based on what is known from the Bible but giving more detail of a mainly sentimental nature.

The OS translation was compiled in Vadstena, at the end of the 14th century (Klemming SFSS 15:iii). The oldest manuscript is Codex Bergmanianius, dated at ca 1420 (ibid.).

Bo is a relatively long text, over 68000 words long. A frequency search was made on chapters 9-17 (pages 1-54 in the SFSS edition), a total of ca 14300 words.

- \textbf{Period III (1450-1526)}

\textit{Prosaiska krönikan (PK)}

PK is an account of the history of Sweden until the reign of Karl Knutsson. It was most probably written down ca 1452, no later than 1457 (Geete 1903:119) and preserved in a manuscript that can be dated to ca 1460 (ibid., p. 118). The complete text, about 4700 words, has been studied.
Didrikssagan (Did)

The Swedish translation of the story of Didrik from Bern is dated at 1450 (Henning 1970:28), but the translation is based on a much older (ca 1250) translation into Old Norwegian.

The text is almost 54000 words long. For the purpose of the present study chapters 1-44 were chosen (ca 10945 words).

*Själen’s kloster* (Linc JB) and *Legenden om Sankt Amalberga* (Linc NR)

In Period III two translators were very active, namely Jöns Budde (JB) and Nils Ragvaldi (NR). Both were born ca 1440 and translated religious prose but most likely came from different parts of the country: JB from Österbotten and NR from Östergötland (Hirvonen 1987:59).

For the present study, two texts were chosen, both preserved in the so-called *Linköping legendariet* (Linc) from ca 1520. The first is *Själen’s kloster* ’the monastery of the soul’, translated by JB ca 1480 (Linc JB), the other is the legend of St. Amalberga (Amalberga of Temse) in translation by NR from the same time (Linc NR).

From Linc JB pages 29-75 were chosen (ca 11088 words) and Linc NR was excerpted as a whole (9835 words).

*Historia Trojana* (Troj)

Historia Trojana is based on the Latin version of the history of Troy by Guido de Columnas (Hirvonen 1987:63). It is included in the study, even though being written down in 1529 it formally belongs to the older New Swedish period (which commences in 1526). In line with a long tradition it was regarded in this study as the last Old Swedish text (Ståhle 1967:121, Hirvonen 1987:63, Delsing 1999, Håkansson 2008:21).

The whole text is over 44000 words long and chapters 1-39 (ca 9780 words) were chosen for the present study.

1.4. The method

Grammaticalizing items show a tendency to rise in frequency, particularly at inception (Traugott 2011:28). Therefore, frequencies of the demonstrative and the numeral will be a subject of study in the empirical chapters.

Historically, the demonstrative *(h)iinn* cliticizes onto the noun before any texts were written down. Thus it is already a clitic -IN in period I (1225-1375). Its
functional range is however limited in comparison to the MS system (see 1.5). The spread of -IN to encompass new functions is studied in a twofold way:

1. a general frequence study—number of occurrences of -IN per 1000 words.
2. a study of frequency of particular functions: out of the first 100 definite NPs, how many are used e.g. anaphorically (see 1.5 and Chapter 2 for a description of different uses of -IN).

Apart from the overall frequency, frequency in particular contexts is also studied; the contexts are presented in Chapter 2. As the frequency of -IN gradually rises in the periods studied, one finds many instances in very short text fragments, while the older texts need to be searched in extenso to yield a handful of examples. Therefore the detailed study is limited to the first 100 occurrences of -IN in a given text. The excerption is purely manual, since it relies heavily on the context. Examples with -IN are sorted according to their:
- form (number, gender, case of the noun)
- function (anaphora, indirect anaphora, unique reference, generic reference)
- function in sentence (subject, object, attribute, prepositional object)
These three types of sorting allow generalizations as to:
- where in the sentence -IN occurs
- in what function it is used
- which nouns take -IN

The numeral *en* ‘one’ is used only as a numeral in the oldest sources. Its grammaticalization is therefore entirely accessible in the OS texts. The spread of the numeral through new functions is also studied in a twofold way.

1. a general frequence study—number of adnominal occurrences of EN per 1000 words.
2. a study of particular functions—how many uses are numeral, specific or nonspecific (see 1.5 and Chapter 2).

With the exception of Sthlm JB all texts chosen for the study are available in electronic form as well as in ‘paper’ editions. This makes it possible to search the text corpus by means of concordance-creating programs, such as KWIC. It is very practical with the numeral *en* ‘one’, as just two search strings will return all\(^6\) forms of its paradigm\(^7\) (compare tables a and b).

---
\(^6\) A number of words beginning with either *en* or *et*, e.g. *ensampnin* ‘alone’ will also find their way to the concordance, which must be manually pruned.
\(^7\) Occasionally, the numeral is spelled *aen*.
Demonstrative *(h)inn* is an already cliticized form -IN in OS. Therefore the search for the marked nouns must be manual. Thus some of the longer texts could not be searched in their entirety; instead fragments of 10000-15000 words were chosen.

*Smaller statistical studies*

In some cases it was worthwhile to look at a particularly frequent noun, e.g. *bonde* ‘yeoman’ in OS legal texts, *biskop* ‘bishop’ in religious prose or *konung* ‘king’ in profane narratives, as well as unique referents such as sun, moon, world, earth. For these the concordance creator KWIC is employed.

### 1.5. (In)definiteness in Modern Swedish

#### 1.5.1. NP types in MS

Definiteness in MS is marked by postposed definite article and preposed indefinite article (only in the singular, though pronoun *några* ‘some’ is used to some extent in this function).

(1a) Jag har köpt *en bil*.
I have bought a car

(1b) Jag har köpt *bil-en*.
I have bought car-DEF
Further, there is a preposed definite determiner which co-occurs with the definite article in definite NPs with an attribute:

\[(2a) \text{Jag har köpt} \quad \text{den svart-a bil-en.} \]
I have bought DET black-W car-DEF

\[(2b) \text{Jag har köpt} \quad \text{de svart-a bil-ar-na.} \]
I have bought DET black-W.PL car-PL-DEF

The preposed determiner can be omitted in lexicalized phrases, similar to proper names:

\[(3a) \text{Gula floden} \quad \text{yellow-W river-DEF} \]
‘Yellow River’

\[(3b) \text{den gula floden} \quad \text{det yellow-W river-DEF} \]
‘the yellow river’

Bare nouns also appear frequently in Swedish, often with amorphous reference (Swedish *dividua*). In SAG (vol. 3:43-44), all NPs that are not overtly definite are regarded indefinite\(^8\). In other words, the overt contrast is definite : indefinite/unmarked. However, in a different chapter, SAG shows that the bare nouns are not always a shortcut to an indefinite article:

\[(4a) \text{skaffa sig körkort} \quad \text{–} \quad \text{skaffa sig ett körkort} \quad \text{get REFL driving licence} \quad \text{get REFL a driving licence} \]

\[(4b) \text{körkort} \quad \text{driving licence} \]

\[\text{‘get a driving licence’} \]

---

\(^8\) _Strukturellt kännetecknas den indefinita nominalfrasen negativt, dvs. av att inte innehålla någon av definitetsmarkörerna (definita attribut, bestämd form av substantivet eller av adjektiv och adjektiviskt böjda ord). (SAG vol. 3:43)_
(4b) ligga på rygg\^9 – ligga på ryggen
lie on back lie on back-DEF
‘lie on one’s back’

(SAG vol. 3, s. 175)

Both are examples of lexicalized phrases, where the noun regularly appears without an article. A paraphrase reveals whether the missing article be definite or indefinite. ‘Zero’ marking can be a proxy for either definite or indefinite marking.

A further point that should be made in connection with examples (4a) and (4b) is that not only are the nouns neutral with respect to definiteness (even though a paraphrase with either definite or indefinite article may be forced upon them), they are also neutral with respect to number. Consider these phrases with plural subjects:

(5a) Alla artonåringar skaffade sig körkort
all eighteen-year-old-PL got REFL driving licence
denna sommar.
that summer
‘All eighteen-year-olds got driving licence that summer’.

(5b) Alla tjejer låg på rygg och solade
all girl-PL lied on back and sunbathed
‘All girls lied on their backs and sunbathed’.

The definite article in (5b) was a result of interpreting the noun phrase as ‘one’s back’. In a sentence han låg på rygg/ryggen ‘he lied on back/the back’ it would be the same as han låg på sin rygg ‘he lied on his back’. But the form of the noun remains the same even if more than one back is involved—and the sentence alla tjejer låg på ryggor/ryggarna ‘all girls lied on backs/the backs’ is ungrammatical. It seems that the language does not handle such phrases in the plural very well, although the plural form is acceptable for some phrases:

(6a) barnets ålder – barnens åldrar/ålder
child-DEF-GEN age child.PL-DEF-GEN age-PL/age
‘the child’s age’ ‘the children’s age’

(6b) Det här tyget ligger på olika bredder/bredd.
this here fabric lies on different breadth-PL/breadth
‘This fabric differs in breadth from one place to another’.

(SAG vol. 2cd:21)

\^9 Different body parts appear in different forms, however. It is not possible to use bare noun in ligga på mage, only the definite form is acceptable, ligga på magen ‘lie on the stomach’.
Bare nouns appear also by names of professions where English requires an indefinite article, as in:

(7a)  Jan är forskare.
     Jan is scientist
     ‘Jan is a scientist’.

(7b)  Jan är präst.
     Jan is priest
     ‘Jan is a priest’.

The Swedish phrase lacks the indefinite article that the English phrase has. The article appears when the name of the profession is accompanied by an attribute and the NP has an attributive character.

(8a)  Jan är en fantastilös forskare.
     Jan is an unimaginative scientist
(8b)  Jan är en utbränd präst.
     Jan is a burnt out priest

Swedish uses an article only if an attribute is present (see example above). The NP forskare ‘scientist’ has an attributive function (Strawson 1950, Burton-Roberts 1976:428). It is not paraphrasable by ‘a certain’, as it is in other structures:

(9a)   A (certain) man was lurking in the garden.
(9b)   John is a (*certain) scientist.
(9c)   *Jan är en viss forskare.
     Jan is a certain scientist.

The argument becomes all the more clear when one considers the following correct sentences, where references are made to professions, also through the attributes (which are different from examples (8a) and (8b), because ‘unimaginative scientist’ is not a profession):

(9d)   Man ska säga att han är katolsk präst.
     One should say that he is Catholic priest.
     ‘One should say that he is a Catholic priest’.
(9e)   År 1880 blev han legitimerad francisk läkare i Paris.
     Year 1880 became he registered French doctor in Paris.
     ‘In 1880 he became a registered French doctor in Paris’.

(all examples from SAG, vol. 3, s. 177)
Consider also the difference between the bare noun and the indefinite article in the following examples:

(10a) Han är bödel.
he is executioner
‘He is an executioner.’

(10b) Han är en bödel.\textsuperscript{10}
he is an executioner
‘He is a tormentor/tyrant.’

In (10a) it is the profession of the subject that is meant, whereas in (10b) the noun is used metaphorically.

1.5.2. The functions of the definite article

The definite article (suffix) in MS is used roughly like the definite article in English. It can be used deictically, to refer to objects in the speaker’s and/or hearer’s immediate presence, as in (11a) and (11b).

(11a) Akta bordet!
Mind table-DEF
‘Mind the table!’

(11b) Kan du ge mig boken?
can you give me book-DEF
‘Can you give me the book?’

It may also be used with referents that are immediately accessible within the discourse (anaphora), directly as in (12), or indirectly via a different discourse referent, as in (13).

‘Once there was a king who had a daughter. The daughter loved to sing.’

\textsuperscript{10} There is a further difference between the two examples, which is only made in speech:

Han är en riktig bödel.
he is a real executioner
In the first sentence, again the profession is meant, whereas in the second it is the metaphorical sense that is underlined.
(13) Han har skaffat en ny bil. Bagageluckan är rymlig.
‘He has bought a new car. The boot is spacious.’

In both deictic and direct anaphoric uses the definite article may be exchanged by a demonstrative, *denna* or *den* ‘this’. This is not possible in indirect anaphora.

(14a) Kan du flytta på detta bord?
‘Can you move this table?’

(14b) Han har skaffat en ny bil. *Denna bagagelucka är rymlig.
‘He has bought a new car. *This boot is spacious.’

In contrast to English, MS uses the definite suffix with inalienables such as body parts where English uses possessives.

(15) Han höll händerna i fickorna.
‘He kept his hands in his pockets’.

The definite article is also used with unique referents, such as the Sun, or, more locally, the king, the bishop etc. It may also be used to mark generic reference. However, in Modern Swedish generic reference can be expressed by a variety of forms.

(16a) **Lejon** är ett däggdjur.
lion is a mammal

(16b) **Lejon-et** är ett däggdjur.
lion-DEF is a mammal

(16c) Ett **lejon** är ett däggdjur.
a lion is a mammal

(16d) **Lejon** är däggdjur.
lions.PL are mammal.PL

(16e) **Mammut-ar-na** var ofantligt stora däggdjur.¹¹
mammoth-PL-DEF were enormously large mammal.PL

---

¹¹ Additional notes:

A similar example can be found in Noreen 1904 (Hästarna äro snabbfotade djur ‘The horses are quick animals’, see Pettersson 1976:132); Pettersson remarks that plural definite generics are highly unusual in Swedish (Pettersson 1976:132).
There are thus five different ways of expressing generic reference in Swedish: three in the singular (bare noun, definite and indefinite) and two in the plural (bare noun and definite). The sixth, indefinite plural, if we take några ‘some’ to be a plural counterpart of the indefinite en/ett (see example (40c) above), is ruled out by definition as it literally refers to ‘some of’, or ‘a certain amount of’ whereas the nature of generic reference is that it refers to ‘all of’. These five structures are not fully synonymous.

The indefinite singular is only used when the description fits any arbitrarily chosen member of the class, that is, it cannot refer to a quality of the class as a whole. Therefore it is incorrect to say:

(17a) *En skrivmaskin uppfanns inte i Sverige.
‘A typewriter was not invented in Sweden’.

Here, the definite form is necessary:

(17b) Skrivmaskinen uppfanns inte i Sverige.
‘The typewriter was not invented in Sweden’.

That is because the definite generic can refer to the class as a whole with a collective meaning. This is the meaning of inclusiveness which we recognize from Hawkins 1978. The contrast between the definite and indefinite singular is one between collective and distributive reference. The definite article may be used generically when reference is made to a feature typical for the whole class, whereas the indefinite requires that the feature apply to each and every member of the class (Pettersson 1976:124, SAG vol. 3:174).

(18a) Den svenske socialdemokraten lever och dör för sitt parti.
‘The Swedish social democrat lives and dies for his party’.

---

Similarly in English by different types of generics: one that predicts something about each member of a given class (‘individual generics’) and one that predicts something about the class as a whole.

a. The squid lives on seaweed.
b. Squids live on seaweed.
c. A squid lives on seaweed.

a. The dodo is extinct.
b. Dodos are extinct.
c. *A dodo is extinct.

(C. Lyons 1999:181-4)
En svensk socialdemokrat lever och dör för sitt parti.

A Swedish social democrat lives and dies for his party.

After Guillaume (ref. Pettersson 1976:125) Pettersson regards the reference in (18b) as more specific, e.g. in a context where a particular member of the party has been caught thinking more of himself than the party.

Both the definite and the indefinite generics contrast with the bare noun generics. The article-less form is used in generic utterances and lacks a plural form—it is neutral not only with respect to definiteness, but also with respect to number.

Gärdsmyg är en flyttfågel.
Wren is a migratory bird

En gärdsmyg är en flyttfågel.
A wren is a migratory bird

Gärdsmygen är en flyttfågel.
Wren is a migratory bird

The contrast between the indefinite generics and the bare noun generics is not well-studied and difficult to grasp. Whereas the indefinite applies to each and every member of a given class, the bare noun present ‘the platonic ideal’ of a member of this class (Pettersson 1976:124).

1.5.3. The functions of the indefinite article in MS

The indefinite article in MS is used when the speaker introduces a new referent to the discourse either assuming that the hearer is unfamiliar with the referent or is himself unfamiliar with it.

Det finns en bil jag gärna vill köpa.
A car I willingly want to buy

Jag har köpt en bil.
I have bought a car
In the so-called negative polarity contexts, e.g. in the scope of negation and in questions, någon ‘some, any’ is used rather than the indefinite article:

\[
\text{(21a) Han har inte gjort någon kaninbur.}
\]

‘He has not made a rabbit cage’.

(Nivre 2002:8)

The use of the indefinite article is not incorrect, though its use renders a different interpretation:

\[
\text{(22a) Han har inte gjort en kaninbur, han har gjort en fågelbur.}
\]

‘He has not made a rabbit cage, he has made a bird cage’.

\[
\text{(22b) Han har inte gjort en kaninbur, han har gjort många.}
\]

‘He has not made one rabbit cage, he has made many’.

In (22a) the classification of the product rather than its existence is questioned and in (22b) the cardinality (EN is thus treated as a numeral rather than an article).

In questions någon is the unmarked alternative (23a), unless the speaker expects a positive answer (23b), see SAG and Nivre 2002:10.

\[
\text{(23a) Har du köpt några kläder på rea?}
\]

‘Have you bought any clothes on sale?’

\[
\text{(23b) Har du redan köpt en julgran?}
\]

‘Have you already bought a Christmas tree?’
INTRODUCTION

Other negative polarity contexts include hypothetical conditionals, comparisons and so-called verbs of volition (e.g. ‘want’). Någon can also appear in contexts other than negative polarity instead of the indefinite article but with a different function.

(24a) Jag talade med någon medicinsk expert.
I spoke with NGN medical expert
‘I spoke to some medical expert’

(Nivre 2002:12)

In (24a) more precise information on the expert is either unavailable or considered irrelevant.

(24b) Vi fortsatte någon kilometer genom skogen.
We continued NGN kilometer through wood-DEF
‘We continued one or two kilometers through the wood’.

(Nivre 2002:17)

In (24b) någon is an approximation of one (one or two, about one, hardly more than two, SAG: 416).

The Swedish någon thus has three fairly distinct uses, in which it contrasts with the indefinite article en:

1. in negative polarity contexts it is the unmarked determiner in noun phrases with indefinite reference;
2. in other contexts it is the marked alternative to the indefinite article in noun phrases with indefinite reference, usually indicating that the speaker is unable or unwilling to give further information about the referent;
3. in quantative noun phrases någon can perhaps be described as an ‘approximate numeral’, meaning ‘about one’ and contrasting with the indefinite article in its numeral sense of ‘(exactly) one’. (Nivre 2002)

1.5.4. Use of bare nouns in MS

MS nouns may be singular or plural and they are classified into neutrum (neuter gender) and utrum (non-neuter, a coalescence of the OS masculine and feminine genders). They are also divided into count (individua, Lat. ‘which cannot be divided’) and mass (dividua, Lat. ‘which can be divided’) nouns and this division is important for article usage in MS.

Generally, the presence or absence of an article (either definite or indefinite) will influence the interpretation of the noun as either count or mass. Consider the following MS examples:
(25a) Tydliga spår både av älg och hund fanns i den nygrusade vägbanan. 
‘There were clear traces of both elk and dog in the newly groveled road.’

(25b) Den lukta både får och get. 
‘It smelled of both sheep and goat’.

(25c) Vi såg spår av tre älgar. 
‘We saw traces of three elks’.

(25d) Vi såg en stor älg. Den stora älgen tittade på oss. 
‘We saw a large elk. The large elk stared at us’.

The lack of articles or other determiners/quantifiers in example (25a) renders älg ‘elk’ uncountable; what the observer saw was älgspår ‘elktraces’, without commitment to how many elks may have left them. It is possible it was just one, it is equally possible there were a hundred. In (25c) ‘elk’ is countable, by means of a numeral three. In (25d) the usage of either article again renders the elk countable.

The opposite is possible as well: by adding an article to a noun with a typical mass reading, the speaker makes it countable (and ‘divisionable’), as in (26b) and (26c) below:

(26a) Jag gillar vin. 
I like wine 
‘I like wine’.

(26b) Jag gillade vinet. 
I liked wine-DEF 
‘I liked the wine’.

(26c) Det var ett gott vin. 
It was a good wine 
‘It was a good wine’.

The bare noun (as älg ‘elk’ in (25a) above) is neutral with respect to definite-indefinite opposition—it is not definite and it is not indefinite. It is further neutral with respect to number: it is neither singular nor plural. By using a bare noun, the speaker may further withdraw gender classification, consider (27b)
(27a) Han ska köpa bil.
he will buy car
‘He will buy a car’.

(27b) Bil är dyrt.
car is expensive-NEUT
‘It is expensive to buy a car’.

(Delsing 1997:58-59)

1.6. Organisation of the book

This book is organised as follows. In Chapter 2 some theoretical assumptions about grammaticalization, in particular grammaticalization of articles, are presented. Chapter 3 presents the etymologies of the articles alongside reconstructions of the cliticization of (h)inn, which predates the OS sources. Chapters 4 and 5 contain a detailed overview of the development in consecutive periods, of definite and indefinite articles respectively. In Chapter 6 all the results are brought together and the interdependencies between the grammaticalization of definite and indefinite article analyzed.

Note that not all nouns can be used bare here, it seems the act of buying must be more than trivial:

*Han ska köpa bok.
he will buy book
‘He will buy a book’.

*Bok är dyrt.
book is expensive
‘It is expensive to buy a book’.

(Delsing 1997:58-59)
2 Grammaticalization and (in)definiteness

2.1. Introductory remarks

Grammaticalization has played a vital role in historical linguistic studies of the last thirty years, with a large number of papers, monographs and conferences devoted to the phenomenon. Discussion has been both empirical, with large studies focusing on its occurrences in different languages, and theoretical, with numerous attempts to define the concept and its ingredients. Detailed accounts of both can be found in, among others, Campbell and Janda (2000) and the latest and most exhaustive overview is given in Narrog and Heine (2011) (The Oxford Handbook of Grammaticalization).

The main focus of the present chapter is limited to the relationship between grammaticalization and definiteness. However, some introduction is necessary. Therefore, the chapter is organized as follows: first, grammaticalization is defined and its hallmarks discussed, together with some examples of the phenomenon. A special note is made of the role of context in grammaticalization and the types of contexts. Further, the concept of grammaticalization chains is presented. Two relevant chains—the development of the definite article out of a deictic element and of the indefinite article out of numeral ‘one’ are presented in more detail. Finally, the possible correlation between these chains is discussed.

2.2. What is grammaticalization?

The term grammaticalization was coined by A. Meillet in 1912, but the phenomenon to which it refers had already been observed and described in earlier studies, e.g. by von der Gabelentz.

Was heute Affixe sind, das waren einst selbständige Wörter, die nachmals durch mechanische und seelische Vorgänge in dienende Stellung hinabgedrückt wurden. (von der Gabelentz 1901:255)
For Meillet grammaticalization was a process whereby autonomous lexical items gained grammatical properties (l’attribution du caractère grammaticale à un mot jadis autonome, Meillet 1926 [1912]: 131) and in doing so changed the system.

La ‘grammaticalisation’ de certains mots crée des formes neuves, introduit des catégories qui n’avaient pas d’expression linguistique, transforme l’ensemble du système. (Meillet 1926 [1912]: 133)

A wider definition comes from Kuryłowicz, who observes that grammatical items may likewise undergo grammaticalization.

Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivative formant to an inflectional one. (Kuryłowicz 1975:52)

Later authors expand on this definition, rephrasing it or shifting the focus (for more definitions see Heine and Reh 1984:15, Hopper and Traugott 2003 [1993]:18, Lehmann 2004:155, Haspelmath 2004:26, Narrog and Heine 2011:2f.), though in essence it remains the same. There is some disagreement concerning where grammaticalization starts and ends: some authors distinguish the change of a lexical item into a grammatical one from the change of a grammatical item into another—e.g. Traugott (2002:26f.), who terms the former ‘primary’ and the latter ‘secondary’ grammaticalization. The formation of the definite and indefinite articles belongs to the second group.

Two classical and widely discussed examples of grammaticalization include:
1. the development of the French pas ‘step’ into a negation marker (Meillet 1912);
2. the development English to be going to into a future auxiliary (e.g. Fischer and Rosenbach 2000:3).

Pas and going to illustrate two different types of change: from a lexical to a grammatical item and from a grammatical item to a more grammatical one.

The French lexeme pas ‘step’ has given rise to a negation construction ne .. pas. Originally it was one of several elements that could be used to emphasise the negation, comparable with the English ‘not a step further’. As such it was used alongside verbs of motion, preferably verbs denoting motion measurable in steps. As its use gradually spread through other types of verbs, so has pas been reinterpreted as part of the negation rather than an emphatic marker.

The English going to develops from a literal description of an activity (motion) to a future marker. The gradual development from less to more grammaticalized meaning is illustrated by the following examples.

\[14\] With respect to the phenomenon. In the sense of view of grammar (e.g. Hopper) obviously the definition of ‘grammaticalization’ is wider. This is not how the term is used here.
What is grammaticalization?

(1a) I am going (to Harlem) to visit my aunt.
(1b) I am going to marry (tomorrow).
(1c) I am going to like it.
(1d) It is going to rain.
(1e) I am going to go there for sure.
(1f) I’m gonna go.

(Fischer and Rosenbach 2003:3)

Both instances of grammaticalization reveal a number of changes on different linguistic levels:

– phonological (*going to* > *gonna*)
– morphosyntactic (*pas* develops a negative meaning and its position is fixed)
– functional (both come to be used outside their original contexts, e.g. *pas* can be used alongside any verb, not just verbs of motion that may be measured in steps though the original meaning was ‘not a step further’)

These changes may be grouped into those mainly concerning the form of the grammaticalizing item (its phonological representation, position, degree of bond-edness) and those concerning its function (becoming an auxiliary).

We can further note that we were able to illustrate the grammaticalization of *going to* exclusively using modern examples. Similarly, the lexeme *pas* still exists in French in its original meaning of ‘step’. Grammaticalization need not involve the loss of the original form, nor its preservation. The numeral *en* ‘one’ from which the indefinite article in Swedish is descended, is still used, while the demonstrative *hinn* ‘yon’, the source of the postposed Swedish definite article, is lost (apart from the lexicalized phrase *hin håle* ‘the hard = the devil’). The co-existence of different uses of a given form, representing different stages of grammaticalization, implies that grammaticalization is not limited to diachrony but has a synchronic dimension as well; the most radical interpretation of this is Hopper’s concept of emerging grammars (Hopper 1987) which entirely shifts the focus on to synchrony. Although we may order the constructions diachronically (from *I am going* in the purely physical sense to *I’m gonna go* with the future meaning and reduced form), we also find that they co-exist in language (Hopper 1991:22).

Grammaticalizations of lexical items are probably less common and certainly less regular cross-linguistically (how many languages have developed their negation markers from the lexeme ‘step’?). While ‘step’ does not usually develop into negation marker, constructions similar to English *going to* have developed into future markers in a number of other languages (see Heine and Kuteva 2002:161ff). It seems some developments are more common cross-linguistically than others, e.g. demonstratives into definite articles and numeral ‘one’ into an indefinite article. Grammaticalizations of grammatical items show more regularities, allowing the proposition of so-called grammaticalization chains (or paths or channels), which
are similar developments, for unrelated languages. The grammaticalization of both definite and indefinite articles is an example of such changes, since both articles arise out of grammatical forms: the definite out of a demonstrative pronoun and the indefinite out of numeral ‘one’. We will also see that these developments are widely recognized examples of grammaticalization chains. But before taking up those, some major tenets of grammaticalization will be discussed.

2.3. Major tenets of grammaticalization

The passage from a less to a more grammatical item is characterized by a number of phenomena, some examples of which we saw above. These can be grouped into changes concerning the form of the grammaticalizing item and changes concerning its function and use. Lately, a further element, the context, has become the focus of attention (Heine 2002, Diewald 2002). An important claim much discussed in recent literature, is the unidirectionality hypothesis.

2.3.1. Changes in meaning

Semantic bleaching

In the course of change from lexical to grammatical the original lexical meaning, or some part of it, is lost. The paradigm case of this is the fate of the French pas, which, as a negation, has nothing in common with the lexeme pas.

It is less clear what form of semantic bleaching is found in cases of further grammaticalization of an already grammatical item. As the subject of this book is the grammaticalization of articles, let us look at some examples of such developments in different languages.

The source of the definite article is most commonly a demonstrative pronoun (see 2.5.1). In many languages, however, there is more than one. The semantic difference between them is typically that of increasingly precise information about the location of the referent in respect to the speaker or hearer or both. Classical Armenian allowed a threefold distinction:

<table>
<thead>
<tr>
<th>Demonstrative</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ays</td>
<td>‘this (near me)’</td>
</tr>
<tr>
<td>ayd</td>
<td>‘that (near you)’</td>
</tr>
<tr>
<td>ayn</td>
<td>‘that (near him etc.)’</td>
</tr>
</tbody>
</table>

(after Lyons 1999:55)

15 Other frequently employed terms are semantic reduction/loss/attrition/fading/decay, depletion/impoverishment, weakening, desemanticization, generalization of semantic content, abstraction (Campbell and Janda 2000:118).
As the definite article(s) developed from these demonstratives, they retained some of the spatial reference while gaining the definite one.

<table>
<thead>
<tr>
<th>Definite article</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>-s</td>
<td>‘the (near me)’</td>
</tr>
<tr>
<td>-d</td>
<td>‘the (near you)’</td>
</tr>
<tr>
<td>-n</td>
<td>‘the (near him etc.)’</td>
</tr>
</tbody>
</table>

(Definite article Table: after Lyons 1999:55)

However, the Classical Armenian system does not survive in the modern language; instead, the -s and -t forms were reinterpreted as possessives and the only definite article (and third person possessive) is the -n/-d, descended from the classical -n (Feydit 1969, Lyons 1999:55). It seems that the spatial distinctions of the demonstratives are lost along with their grammaticalization into definite articles.

A more complex phenomenon is when the form in question seems to split in two: it undergoes grammaticalization and in doing so loses (some of) its original meaning, but it also survives intact without losing its meaning (a phenomenon sometimes referred to as layering, Hopper 1991:22). The end product seems to be two homonymous forms, as in the French _pas_, but the link between them may still be tangible, as in the English _be going to_.

The term itself, bleaching/loss/depletion etc. leads to the view of grammaticalization as ‘a kind of impoverishment, or deficit’ (Traugott 1988:406) (so does the terminology for the phonetic aspects of the process). This is perhaps misleading, since there is (possibly) loss on the lexical side, but there is also gain in grammatical meaning (Heine et al. 1991:109).

Semantic bleaching has in many works been closely associated with grammaticalization. However, the process as such has been rightly pointed out to be independent of the grammaticalization processes. A number of objections have also been raised regarding its indispensability for grammaticalization. In Traugott (1988) we find a modified stand on the role semantic bleaching plays in grammaticalization:

Bleaching and grammaticalization must be uncoupled (...) certainly, bleaching does occur (...) [but] only in the later stages of grammaticalization (Traugott 1988:407)

This is taken up in her later work with P. Hopper:

There is no doubt that over time, meanings tend to become weakened during the process of grammaticalization. Nevertheless, all the evidence for early stages is that initially there is a redistribution or shift, not a loss, of meaning (Hopper and Traugott 1993:88)

---

16 For more examples of similar developments see Lyons 1999.
Indeed, were we to imagine a grammaticalization process starting with semantic bleaching, we would have to conceive of a form first losing its meaning before gaining a new (grammatical) one. However, there is an affinity in meaning between consecutive forms as the grammaticalization proceeds. Reflexes of the original semantics are present in definite article as developed from a demonstrative pronoun, or a future marker formed out the verbs ‘to wish’ or ‘to want’17. In the same vein, Heine (2002) claims in his discussion of the different contexts in which a form may appear in different stages of grammaticalization, that there are two meanings available for the form in the stages of grammaticalization which he terms bridging and switch contexts (see 2.4): as the source and the target (Heine 2002:85)18.

This may help us better understand why the Classical Armenian definite articles retained the spatial reference of the demonstratives, but also why this reference was lost with time.

On the other hand, the grammaticalization of the indefinite article constitutes a more complex case: while the singularity of the referent is still asserted, new, additional information is captured in the original numeral: that of presenting a new referent. In this sense the original singularity, though suppressed, does not disappear entirely, whereas a new meaning is added. This shows that the concept of bleaching is not straightforward.

2.3.2. Changes in form

Phonological attrition19

As was the case with semantic bleaching, authors seem to agree, at least superficially, that phonological attrition is inherently connected with grammaticalization, though only some consider it a diagnostic of grammaticalization, e.g. for Heine attrition is predicted by grammaticalization theory (Heine 1993:106).

17 A radically different view is presented by Haspelmath: it is not so much that semantic bleaching and phonological reduction go hand in hand, but semantic generalization is in a sense the cause of the other processes of grammaticalization (Haspelmath 1999:1062). This is taken up by Campbell and Janda: [t]hrough a different route, Haspelmath has arrived at the same view as that of detractors of grammaticalization, namely, that the semantic change is not a crucial part of the definition of grammaticalization and can have effect even when no grammaticalization is involved at all. Given this independence, semantic loss is not diagnostic for grammaticalization (Campbell and Janda 2000:121). That is probably an over-interpretation, as there is no claim about semantic bleaching and grammaticalization being disjoined, rather about the order being reversed as against the model proposed by Traugott and Hopper.

18 It must be noted, as Heine himself does, that his model assumes a more clear-cut division between meaning and context than we usually find in language.

19 Other frequently employed terms are: phonological reduction/weakening, phonetic loss, erosion, Abnutzung (Campbell and Janda 2000:121).
There are numerous instances of forms which lose their phonological content in the process of grammaticalization: *going to* > *gonna*, *full* > *f'd* etc. However, cases where the grammaticalized form retains its original phonological value and phonetic representation are not infrequent (e.g. German auxiliaries *haben* and *sein* (Heine 1994:109)). It may therefore be concluded, as Lessau does among others, that phonological attrition is *neither a necessary nor a sufficient property of grammaticalization* (Lessau 1994:263). The reduction follows any other phonological reduction, also independent of grammaticalization: unstressed forms tend to be reduced. Therefore the auxiliaries, maintaining a relative independence, typically continue as non-reduced forms.

The main driving force behind the reduction in phonology is the frequency of use. This phenomenon has been observed by Zipf, who formulated a statistically based rule which states that frequently used forms tend to be shortened (Zipf 1949). Such a view is generally adopted by grammaticalists:

As the meaning generalizes and the range of uses widens, the frequency increases and this leads automatically to phonological reduction and perhaps fusion (*Parallel reduction hypothesis*, Bybee and Pagliuca 1985:76).

[The higher frequency of use of grammatical morphemes favours what von der Gabelentz ([1891]1901) has called the *Abnutzung* (abrasion) of their phonetic substance (Heine et al. 1991:214).

The Swedish definite article is a paramount example of phonological reduction—the original demonstrative was */i:n:/ in its masculine form and */i:n/ in its feminine form, with either a long consonant or a long vowel. In MS (where the masculine and feminine genders have coalesced), its form is the reduced */in:/ For the indefinite article, it is never stressed whereas the numeral may, but need not be, stressed.

*Loss of syntactic freedom*

The core tenet of grammaticalization is the change in the distributional status of the grammaticalizing form. An item that was originally lexical, gains grammatical meaning which leads to a change in the type of contexts in which it may appear. Also a grammatical item may gain additional grammatical meaning further limiting its distribution. A demonstrative may have a relatively free position against the noun, with preference for preposition, postposition, or both, while an article has a fixed location. With the grammaticalization of an originally lexical or less grammatical item, its distribution is changed. The change may depend on the change in the status of the emerging form: from a lexeme to a clitic or from a clitic to an affix and further an inflectional ending. But even when a form remains relatively free, and does not cliticize, its distribution might become constrained. Take
for instance the English definite article. Although it does not attach to either the
noun or any other constituent of the NP, its position is obligatorily initial in the
NP; thus the order *great the feat is ungrammatical.

The change in the level of bondedness of the form is presented on the so-
called cline of grammaticalization

\[ \text{content word} \rightarrow \text{grammatical word} \rightarrow \text{clitic} \rightarrow \text{inflectional affix} \]

(Hopper & Traugott 1993:7)

It has been suggested that as the first change is different from the following
ones since it involves a move from an open to a closed word class, the cline should
be split in two:

\[ \text{content word} \rightarrow \text{grammatical word} \]
\[ \text{grammatical word} \rightarrow \text{clitic} \rightarrow \text{inflectional affix} \]

(e.g. Faarlund 2007)

The cline of bondedness of the form is a simplification. The clitics, particu-
larly, form a heterogeneous group—some show cliticization gaps and will only
attach to certain hosts. Some affect the phonology of the host or are affected them-
selves. The much cited English genitive -s does not have as unrestricted a distri-
bution as it is claimed to have: it will not attach to plural nouns where the plural
is marked by -s, while it freely attaches to singular nouns ending in -s (see also
Börjars 2003).

As mentioned above, the first to formulate the cline of grammaticalization
was Meillet, but the cline has been both reworked and extended by several au-
thors (independently) since. These additions and changes have often fallen prey to
a confusion of the different levels of analysis—e.g. in Hopper and Traugott 1993
we find the following cline illustrating the development of the final s in let’s:

\[ \text{word} \rightarrow \text{affix} \rightarrow \text{phoneme} \]

(Hopper and Traugott 1993:13)

Though the analysis of the s may be correct (it is no longer a separate mor-
pheme but has become merely a *phonemic constituent of a (monomorphemic)*
word (Hopper and Traugott 1993:13)), the analysis implies that an affix is ne-
cessarily more than one phoneme, which is easily disproven by the genitive or
plural s in English. In other words, the first two items of the cline illustrate loss
of syntactic freedom, whereas the last two loss of the semantic or grammatical
content.

\[20 \text{ For a detailed analysis of what is and what is not a clitic and what types of clitics there are}
\text{the reader is referred to the account of Zwicky and Pullum (1983) and Anderson (2005).} \]
It is also to be noted that some traces of earlier meanings of an item are likely to survive as constraints on the distribution of the item in its new function, a principle termed persistence.

2.4. The role of context in grammaticalization

The process of grammaticalization has been described in the form of a cline which focuses mainly on the level of (in)dependence of the form in question. Although the authors using this model mention that the meaning of the item as it progresses along this cline, changes towards more grammatical and less lexical, lining up the two changes is not always possible. The fact that a clitic has become an affix does not imply any change in its meaning, nor does the fact that a clitic remains so without becoming an affix deprive it of such a change.

A proposition based entirely on the other side of the coin, namely the semantic change, is the description of the types of contexts (Heine 2002, Diewald 2002; both propositions have much in common and have appeared in the same volume, but are, to the author’s knowledge, independent of each other).

Heine’s proposition differentiates between bridging contexts, switch contexts and conventionalization:

- **bridging contexts**: trigger an inferential mechanism to the effect that, rather than the source meaning, there is another meaning, the target meaning; it is most likely to be inferred but is cancellable (Grice 1967) which means that an interpretation in terms of the source meaning cannot be ruled out; a given linguistic form may be associated with a number of contexts.
- **switch contexts**: incompatible or in conflict with some salient property of the source meaning; interpretation in terms of the source meaning is ruled out; the target meaning provides the only possible interpretation.
- **conventionalization**: meanings that need no contextual support, turn into ‘normal’, ‘inherent’, ‘usual’ or ‘semantic’ meanings. That a target meaning has been conventionalized may be seen when it can be used in new contexts, other than bridging or switch, may violate the source semantics.

(Heine 2002: 84-85)

‘Source meaning’ in his terminology is the meaning of the element prior to the onset of the grammaticalization process.

The notion of ‘bridging context’ as introduced by Heine shares some properties with the ‘bridging inference’ as defined in psycholinguistic literature (e.g. Clark and Haviland 1977, cf. Fraurud 1990:396). ‘Bridging inference’ is the processing of first-mention definite NPs, potentially time-consuming and giving
rise to potential difficulties. In these studies, anaphoric use is seen as the primary function of the definite article. Thus a definite NP lacking an antecedent will always create an interpretation problem. The similarity between the bridging context and bridging inference is conspicuous: in Heine’s model it is in the bridging context that a new inference may be drawn; in other words, the interpretation of a form is not as obvious and natural as in other, more canonical contexts.

Therefore, the rise of a new grammatical meaning can be described as a four-stage scenario:

I. initial stage: context unconstrained, source meaning
II. bridging context: a specific context giving rise to an inference in favour of a new meaning; target meaning foregrounded
III. switch context: a new context incompatible with the source meaning; source meaning backgrounded
IV. conventionalization: the target meaning need no longer be supported by the context, may be used in new contexts; target meaning only.

(after Heine 2002:85)

As Heine himself observes:

The scenario rests on a clear-cut division between context and meaning, although it remains largely unclear how the two interact (Heine 2002:86).

Diewald’s model is similar to Heine’s; she uses the concept of ‘critical context’ rather than ‘switch context’.

2.5. Grammaticalization chains

We have noted before that certain forms show a propensity to grammaticalize into certain other forms, and the tendency is strong even among unrelated languages, forming what are called grammaticalization paths, channels or chains. An extensive presentation of these can be found in Heine and Kuteva 2002.

With respect to the definiteness category the following cross-linguistic grammaticalization chains have been identified:

\[(5) \text{deictic element (typically a demonstrative pronoun)} \rightarrow \text{definite article} \]
\[\text{numeral ‘one’} \rightarrow \text{indefinite article}\]

The renewal of exponents is also testified: there are languages with more than one definite article, all stemming from deictic elements (e.g. Frisian, see Ebert 1971).
In 2.5.1 and 2.5.2 we will take a look at the grammaticalization chains leading to the formation of the definite and indefinite articles respectively. Since the articles always seem to grammaticalize in this particular order: first the definite then the indefinite, we will in 2.5.3 look at a possible way of combining the two chains into one: a grammaticalization of (in)definiteness.

2.5.1. From a demonstrative to a definite article

Definite articles are typically derived from demonstrative pronouns, usually distal, if the language shows differences between proximal and distal demonstratives (e.g. Heine and Kuteva 2002). Other demonstrative elements may also serve as the source of the definite article, e.g. the Sissala article ná has its source in the verb ‘to see’ (Blass 1990, quoted in C. Lyons 1999:331), similar in use to the English verb ‘to behold’ in utterances like *Behold the walls of Moria*, where the verb is used ostensively (corresponding to a gesture or an utterance ‘Here are the walls of Moria’). There have been some attempts to find other sources for the demonstratives, mainly among possessive pronouns, briefly presented below.

There is a certain amount of functional overlap between the definite article and the demonstrative on the one hand and the definite article and the possessive on the other, in probably all article languages. Consider the following examples (after Fraurud 2000b:246):

(6a) Beside the barn there is a little cottage. The/This cottage was built in 1875. (but: *Its cottage..)

(6b) Beside the barn there is a little cottage. The/Its roof is leaking. (but: *This roof is leaking)

Such overlaps may be found in many article languages, although some show a stronger preference for one over the other (e.g. French vs Swedish, as reported in Jonasson 1998 and 2000). In direct anaphoric contexts the demonstrative may always be used instead of the definite article, but this is not allowed in the indirect contexts—for some authors this forms the basis of telling the two apart (e.g. Himmelmann 1997). There are counterexamples, but seemingly only with marked interpretations (Fraurud 2000b:247, compare Kleiber 1990, Apothéloz and Reichler-Béguelin 1999).

The variation between the demonstrative and possessive cases implies that there is no overlap between the two with regard to the definite article. The overlaps between the article and demonstrative concern the direct anaphoric context (see example 6a), and may be seen as an extended deictical use of the demonstrative (Lyons 1975). The overlaps between article and possessive on the other hand mainly
concern what is often termed in literature ‘associative anaphora’ (Ch. Lyons 1999) and what is referred to here as ‘indirect anaphora’ (see example (6b)).

Practically all reconstructions of the grammaticalization of the definite article start with the demonstratives whose use gradually spreads to include even the ‘possessive’ contexts—i.e., indirect anaphora. Thus it is legitimate to ask whether the grammaticalization may start at the other end, namely with a possessive element which gradually enters the demonstrative contexts. Such attempts are presented in Bechert 1993, Frajzyngier 1997 and in Fraurud 2000b, based among others on the extensive use of possessives in many Uralic languages—a use that is often labelled ‘as definite article’ or being ‘instead of definite article’ (see Fraurud 2000b). They are also overwhelmingly bound rather than free forms, showing a certain amount of grammaticalization. They are highly frequent items and their use has partly spread to contexts other than the original indirect anaphora (Fraurud 2000b:250ff). However, the use is not obligatory and Fraurud’s conclusion is that they cannot be considered a grammaticalizing definite article. Similarly, Becher remarks:

... in this continuum (between the possessive suffixes and an emerging definite article—DS), the end seems to be never reached, at the eastern margins of Europe and in Northern Asia, the definite article remains a category in statu nascendi. It might even be the case that this category in Eurasian languages is a product of our Eurocentric perspective. If we reverse the perspective, we can view the European category of the definite article as a special case of a category of belonging which is denoted indirectly, in Europe, by weakened demonstratives and would be more aptly expressed by possessive suffixes, as it is in Northern Asia and its western outskirts. (Bechert 1993:37f.)

We will therefore assume the source of the definite article to be a demonstrative. The first step in grammaticalization must therefore be an extension in use of the demonstrative beyond the purely deictic use. This first extension is constituted by the use of the demonstrative in the anaphoric function, to mark deixis, but in the textual rather than situational context (Lyons 1975). Note that the contrast between the proximal and distal deixis is lost in anaphora: for instance, there is no possibility to vary the use of the demonstratives to point out referents mentioned much earlier and a little earlier in the text. The original distinction into this ‘near me’ and that ‘further away’ is thus blurred.

In this function it is still difficult to see a difference between a demonstrative and an article—in article languages there is however much variation in use between article and demonstrative in anaphora. Although the extension in use constitutes the first step of the grammaticalization process, it is not yet sufficient to make an article.

The second stage in the grammaticalization is the further extension to entities introduced into the discourse indirectly, connected with some items mentioned
earlier—referred to here as indirect anaphora. It is generally agreed that demo-
stratives cannot fulfil this function (however see Charolles 1990 and Apothéloz
and Reichler-Béguelin 1999 for a different description). As mentioned at the be-
inning of this section, in this context the variation is between the article and the
possessive pronoun.

For the indirect anaphora to operate, the first referent need not be definite (see
(6b): a house—the roof). In fact, it need not even be specific (see example (7)).

(7) When a violinist plays a solo, the audience often leaves the room.

Both uses discussed above may be called textual, since they rely on prior in-
formation in the text. The definite article is however used even in contexts where
no such information is provided, also to introduce referents into the discourse. The
article occurs with referents assumed by the speaker to be familiar or identifiable
to the hearer, also on non-textual grounds.

The definite article derived from the demonstrative evolves according to
Greenberg 1978 along the following grammaticalization path:

(8) Stage 0       Stage I       Stage II       Stage III
demonstrative > definite article > specific article > noun marker
(after Greenberg 1978, see also de Mulder and Carlier 2011:525)

The development is more likely to proceed to stage II and beyond in a lan-
guage that has no indefinite article (Greenberg 1991, cf. De Mulder and Carlier
2011:525).

The development from Stage 0 to Stage I can be further subdivided into sub-
stages. It originates with the use of the demonstrative in anaphoric contexts, when
an exophoric (situational) marker is used intralinguistically (e.g. Lyons 1975, Dies-
sel 1999:109-11). A closer examination of the early stages of article formation
reveals that there are some limitations of the anaphoric contexts which may give
rise to the grammaticalization of the definite: these will usually involve a referent
that is an important participant of the narrative but is not the focus of attention (the
data comes mainly from studies on the Romance languages, see de Mulder and
Carlier 2011 for references).

The use of demonstrative in (some) anaphoric contexts is followed by the
spread of the form to contexts that are still textual but indirect, as in (9).

(9) We arrived in a village. The church was on a hill.

The referent of the defNP ‘the church’ must be retrieved indirectly through
the earlier mentioned ‘a village’. This context does not allow demonstratives, al-
though occasional examples to the contrary are cited in literature (e.g. Kleiber 1990, Apothéloz and Reichler-Béguelin 1999). Such examples are at best infrequent and do not constitute a typical demonstrative context. It is therefore the shift from direct to indirect reference that is crucial in the grammaticalization of the article.

De Mulder and Carlier argue that this shift is possible because the source of the definite article is not just a demonstrative but a distal demonstrative (in locative systems) or the 3rd personal demonstrative (in person-oriented systems). Their argument is based on the fact that while a proximal demonstrative allows the hearer to wholly identify the referent through the context, a distal demonstrative is more vague and ‘can be understood as an invitation addressed to the hearer to mobilize previous knowledge in order to retrieve the referent’ (de Mulder and Carlier 2011:530). In this way, the distal demonstrative can denote a referent not fully identifiable and appeal to specific knowledge shared by the speaker and the hearer (see also Himmelmann 2001:833).

As the distal demonstrative in its recognitional use, the definite article can denote a referent that is not accessible in the immediate context of utterance. The distal demonstrative becomes a definite article when the anchorage in the speech situation is lost and the use of the article no longer requires specific knowledge (shared by speaker and hearer) to be activated in order to identify the referent of the noun phrase. Rather, the definite article conveys the instruction that the descriptive content of the NP allows the identification of the referent univocally, by virtue of its structural links with a frame of accessible knowledge, these links often being of a stereotypical nature (de Mulder and Carlier 2011:531). The knowledge may be accessible not through the linguistic context but through the situational one—therefore identification of ‘the king’ will render no difficulties when the phrase is uttered in a monarchy. The situation of the utterance may be more local (identification of the referent of ‘the sofa’ will be limited to the lodgers of the flat where ‘the sofa’ is placed) or less local, ultimately incorporating the whole world (and the unique referents such as ‘the Sun’).

The next stage identified by Greenberg 1978 is of a specific article. The article comes to be used as an indefinite article would: to present new referents in the discourse. Definiteness becomes reinterpreted as specificity (Epstein 2002:368). Himmelmann 1997 suggests that specific articles may evolve directly from demonstratives, without first becoming definite articles (Himmelmann 1997:107). The development would be similar to that of the English demonstrative this (Prince 1981a, 1981b, Diessel 1999:139), as in (10), where it introduces a new, salient referent.

(10) There came this man and demanded to talk to the manager.
Though the referent of ‘this man’ is new in the discourse, it is presented by means of a demonstrative. However, the presentation will only be successful if the referent is picked up again and turns out to be the topic of the story that follows. One difference from the development of the definite article and this in English is that whereas the former involves a distal demonstrative, the latter involves a proximal one.

Greenberg’s model shows the entire grammaticalization of a demonstrative into an article as a single step. A more detailed picture presents the grammaticalization of the definite article in the following stages:

(11) Stage I: the (distal) demonstrative
Stage II: the direct anaphoric marker
Stage III: the indirect anaphoric marker
Stage IV: the uniqueness marker
Stage V: the generalized article

The generalized article here is meant to encompass the generic usage, which according to some authors (in particular Krámský 1972) is the hallmark of a true, full-blown definite article.

In Table 4 below the stages in grammaticalization of the definite article are combined with the types of context discussed above.

Table 4. Stages and contexts in the grammaticalization of the definite article

<table>
<thead>
<tr>
<th>Stage in grammaticalization</th>
<th>Context</th>
<th>Demonstrative/Definite article—function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>Original context</td>
<td>Deixis</td>
</tr>
<tr>
<td>Stage II</td>
<td>Bridging context</td>
<td>Direct anaphora (textual deixis)</td>
</tr>
<tr>
<td>Stage III</td>
<td>Bridging context/Switch context</td>
<td>Indirect anaphora (textual)</td>
</tr>
<tr>
<td>Stage IV</td>
<td>Extended switch context</td>
<td>Unique reference (non-textual)</td>
</tr>
<tr>
<td>Stage V</td>
<td>Conventionalization</td>
<td>Generics (non-textual)</td>
</tr>
</tbody>
</table>

The original context is the (distal) demonstrative pronoun used deictically in a speech situation: when the speaker wishes to draw the listener’s attention to an item in their immediate surroundings but (since it is a distal demonstrative) not close to either of them.

The bridging context constitutes the textual use of the (distal) demonstrative—the text replaces the speech situation, with deixis applied to a linguistic item whose accessibility is grounded in the text and not in the physical situation.
In the account of J. Lyons (1975), the grammaticalization of the definite article originates with the anaphoric use. The bridging context is one where the original meaning is available, while a new meaning is also possible. In this case the original meaning is that of the demonstrative if the conversation is about something accessible through the situational context, and the new meaning is that of—accessibility through the textual context. This can be seen in article languages where the definite article is regularly used anaphorically but where, in many cases, demonstratives are also possible in this usage.

Stage III of the grammaticalization process is an extension of the bridging context: an item is accessible textually but through some other item rather than repetition of an already introduced one. At this stage a situationally motivated reference is still possible. Imagine an electrician coming to your house to examine a faulty kitchen appliance of a strange design. Unable to locate the switch on it he may ask: *Where is the switch?*—here ‘the switch’ will refer to a part of the appliance that has not been mentioned but is present in the situation. In table 4 stage III corresponds to the bridging/switch context. Demonstratives are no longer possible in this stage (see examples (6a) and (6b)). Textually, the extension from the direct anaphora to the indirect may proceed through a use of the form by the referent that is in direct anaphora—but a long distance one and therefore not immediately accessible—which may also be accessed through some other, closer referent. Consider the following OS example:

\[ (12) \text{lægs} \text{ man by woman gets with child it is} \]
\[ \text{frillu} \text{ child marry he later with laws than is} \]
\[ \text{pet} \text{ wife child because that when he bettered} \]
\[ \text{konone} \text{ then bettered he also} \]
\[ \text{barnnit} \text{ child-DEF} \]

‘If a man lies with a woman and gets her with child, it is a bastard. If he marries her later according to the laws than it is a legitimate child, because when he bettered the woman, he also bettered the child’.
The source of the definite marking here is both the direct anaphoric reference to a child, but also the indirect one through the word ‘the woman’. Such contexts give rise to the use of the definite even when the first mention is lacking but the specificity and definiteness of the referent can be inferred from the context. The inference is then based on the general knowledge and stereotypical connections between different concepts, such as marriage: wife, husband, child etc. Both direct and indirect anaphora are textual uses, though the latter also relies on some knowledge outside the text.

The next stage, herein termed the extended switch context, involves the use of the definite by unique referents. It must be noted that this use is clearly at odds with the original demonstrative use where a referent was singled out from among other similar referents, whereas here a referent is marked though it is the only representative of its kind.

Finally, conventionalization of the definite article means that it can be used generically.

### 2.5.2. From numeral ‘one’ to an indefinite article

The source of the indefinite article is not a demonstrative element but the numeral ‘one’, cross-linguistically the only confirmed source of the indefinite. Traditionally, two major uses of the indefinite are proposed: specific and nonspecific.

(13a) I bought a car this morning.
(13b) I am looking for a reliable car.

This double nature of the indefinite article has in literature been termed ‘referential opacity’ (Quine 1953). In some contexts, like (14a), the article is referentially transparent (specific), in others like (14b) it is referentially opaque and requires a wider context to be interpreted as either specific (15) or nonspecific (16).

(14a) John married a rich woman.
(14b) John wanted to marry a rich woman,
(15) ... but she refused him.
(16) ... but he couldn’t find any.
(Quine 1953)

The English indefinite article does not by itself indicate whether the nominal marked by it is intended to refer or not. The situation is similar in MS. But there are languages where this contrast is made. It has also been observed for all Creole
languages that the numeral ‘one’ (the source of the indefinite article) marks referring indefinites, while zero marks non-referring indefinites (Bickerton 1981)\textsuperscript{21}.

According to Heine 1997, the grammaticalization of the indefinite article follows five stages.

Diagram 3. Model of the grammaticalization of the indefinite article according to Heine

<table>
<thead>
<tr>
<th>Stage I:</th>
<th>the numeral ‘one’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage II:</td>
<td>the presentative marker</td>
</tr>
<tr>
<td>Stage III:</td>
<td>the specific marker</td>
</tr>
<tr>
<td>Stage IV:</td>
<td>the nonspecific marker</td>
</tr>
<tr>
<td>Stage V:</td>
<td>the generalized article</td>
</tr>
</tbody>
</table>

(Heine 1997:72-3)

As was the case with the stages of grammaticalization of the definite article, these are not discrete and it is possible for an item to have two (neighbouring) interpretations: such contexts give rise to the new functions. Ambiguity in interpretation between the numeral and the presentative and specific markers in speech is often resolved by the presence or absence of stress.

The presentative marker is in fact a subtype of the specific marker, used to introduce new referents into the discourse. It is reasonable to separate it from the wider specific marker because it seems to be the original non-numeral function of ‘one’ where there is competition from other presentative markers. Hopper and Martin note in their study of the rise of the indefinite article in Old English that there is some variation in this context between the numeral \textit{an} and \textit{sum} and that the variation seems to be quite systematic: while \textit{sum} introduces new and salient referents, with numerous subsequent mentions and is located at the beginning of the text, \textit{an} introduces less salient referents with fewer subsequent mentions and can appear anywhere in the text, not just at the beginning (Hopper and Martin 1987). The presentative marker is thus strictly connected with the textual functions. So the origin of the indefinite article lies, as the origin of the definite, in the textual use.

The specific marker also marks new referents but these need not play any part in the subsequent discourse. It is a marking contrastive to the definite article.

A nonspecific marker is used to indicate that the noun marked is not meant to refer (see example (16)). As has been observed in literature there are a number of contexts where unless the noun is definite it must be regarded as not only in-

\textsuperscript{21} Since the indefinite article in English does no longer differ between referring and non-referring indefinites, one may see the gradual grammaticalization of the proximal this (see 2.5.1) into a referring indefinite article in spoken English as a later stage in the chain of grammaticalization comprising all elements marking (in)definiteness. This renewal is by no means an obligatory part of the development.
definite but also nonspecific (e.g. Barwise and Cooper 1981). These are so-called negative polarity contexts where the noun is found e.g. in the scope of negation, hypothetical conditional, in a comparison or after a verb of volition (e.g. to want). Finally, the generalized indefinite article can be used in generic contexts.

For the sake of a diachronic study, where the speaker’s intentions are difficult to gauge, I will differentiate in Chapter 5 between the presentative, specific and nonspecific uses based on Heim’s concept of the lifespan of a discourse referent, measured by number of further mentions.

The progression of ‘one’ from a numeral to an indefinite article is by Givón 1981 summarized in the form of the following diagram:

Diagram 4. Model of the grammaticalization of the indefinite article according to Givón

quantification > referentiality/connotation > genericity/denotation

One may interpret such a progression as another instance of semantic bleaching along a markedness/implicational space:

(i) Having quantity implies existence/reference.
(ii) Having existence/reference implies having connotation/genericity.

Each of the two transitions above may be viewed as removing one more marked semantic feature of ‘one’; the first step bleaching out quantification, the second bleaching out existence/reference (Givón 1981:50f).

The following table brings together the contexts and the stages of grammaticalization of the indefinite article (both presented by Heine, 2002 and 1997 respectively):

Table 5. Stages and contexts in the grammaticalization of the indefinite article

<table>
<thead>
<tr>
<th>Stage in grammaticalization</th>
<th>Context</th>
<th>Numeral ‘one’/indefinite article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>Original context</td>
<td>Numeral</td>
</tr>
<tr>
<td>Stage II</td>
<td>Bridging context</td>
<td>Presentation marker</td>
</tr>
<tr>
<td>Stage III</td>
<td>Bridging context/switch</td>
<td>Marker of specific reference</td>
</tr>
<tr>
<td>Stage IV</td>
<td>Switch context</td>
<td>Marker of non-specific reference</td>
</tr>
<tr>
<td>Stage V</td>
<td>Extended switch context</td>
<td>Generalized article</td>
</tr>
<tr>
<td></td>
<td>Conventionalization</td>
<td></td>
</tr>
</tbody>
</table>

The original context is the numeral. Its function is to mark that the entity referred to exists and in one exemplar.
The bridging context is the use of the numeral as a presentation marker—the quantity is admittedly still ‘one’, but this information in itself is usually redundant. Rather, a singular referent is marked as being introduced in the discourse. There is a strong likelihood that this referent will be taken up in later discourse. However, a strictly numerical interpretation is still viable.

The switch context involves the use of the form in specific contexts to present new referents important in the story including those that are not ‘persistent’ (Givón 1983).

The extended switch context involves the use of the form by nonspecific referents (in negative polarity contexts and elsewhere). It is impossible to read the original numeral within this, as it assumes non-existence and therefore non-countability.

Conventionalization for the indefinite article implies generic use just as for the definite one. Here the opposition to the original meaning of cardinality is even stronger, since generics refer to the whole class thus directly opposing the notion of there being just one individual answering to the description.

2.5.3. Grammaticalization of (in)definiteness—a large chain?

It has been pointed out in some studies of grammaticalization of (in)definiteness that the grammaticalization of the indefinite is not independent of the grammaticalization of the definite (Stark 2002, de Maulder & Carlier 2011). The grammaticalization of the indefinite article in particular seems dependent on the grammaticalization of the definite having reached at least stage II (direct anaphora) and possibly stage III (indirect anaphora), where the form can no longer be regarded as a demonstrative but rather as a semi-article. Only then does the numeral seem to spread to the presentation contexts. In these a regular difference is made between specific and nonspecific referents—a difference that seems to be more fundamental than the one between the definite and the indefinite. Therefore it is not unusual to find nonspecific presentation markers as well as specific ones; in Swedish *en* is limited to specific uses and in the nonspecific either a bare noun or *någon* ‘some’ is used. The question remains open as to precisely what level of grammaticalization the definite article must reach to set the scene for the grammaticalization of the indefinite one. It seems logical that it must at least be obligatory to mark the prominent referents of the discourse so that the need for properly introducing them is created.

The grammaticalization of the definite article, on the other hand, is also dependent on the existence of an indefinite article—in its later stages the definite article may grammaticalize further into a marker of specific reference and finally into a noun marker (Greenberg 1978). It seems that the presence of an indefinite
article checks this development into a marker of specificity (from Greenberg’s stage I to his stage II).

Cross-linguistically, not only do the grammaticalizations discussed so far (demonstrative < definite article, numeral ‘one’ < indefinite article) occur, but they occur in this particular order. It is therefore reasonable to treat these grammaticalizations as parts of a larger chain of events—the grammaticalization of (in)definiteness.

2.6. Grammaticalization and language change

In grammaticalization studies, language change is regarded as gradual and always involves variation rather than abrupt substitution of one structure by another. A model of language change can be illustrated in the following way (after Norde 2009:16):

\[ A > \{A/B\} > B \]

\{A/B\} symbolizes the variation between old and new forms which for some time coexist until one of them disappears or is reinterpreted. This phenomenon is known as ‘gradience’ (Traugott and Trousdale 2008). The model above is a simplified one, since there are more stages of change conceivable—and all or most may coexist at some point. The model would thus grow into something like the following:

\[ A > \{A/B\} > \{A/B/C\} > \{A/B/C/D\} > \{B/C/D/E\} > \{C/D/E\} \]

From a synchronic point of view, stages such as stage 4, where three different structures (B, C and D) exist in parallel with the original one (A), are problematic in description, since many occurrences will be ambivalent in interpretation between say A and B or B and C. Such is the case of *be going to* in sixteenth and seventeenth century texts (in the Helsinki corpus), where the construction may be interpreted either as motion verb or as future auxiliary (Norde 2009:17, see Brinton and Traugott 2005:26ff).

2.7. Conclusions

Grammaticalization studies belong to the domain of study of language change. They enable a study of how a grammatical category is marked and how that marking changes in time, presenting it in the form of a model of grammaticalization in (non-discrete) stages.
For the category of (in)definiteness a number of studies reveal that the definite articles develop out of demonstrative elements and the indefinite ones out of the numeral ‘one’. These two developments form grammaticalization chains—originating in a common source and progressing toward a common target. Since they concern the same category, they may be reconciled within a larger model where the grammaticalization of the indefinite article is triggered when the definite has become the obligatory marker of the textual functions.
3 Sources of grammaticalization

3.1. Introduction

Before we turn to a detailed study of the grammaticalization of the definite and indefinite articles in the Old Swedish period, 1225-1526 AD, we will consider in this chapter the sources of these grammaticalizations and discuss the developments which predate the OS period. The discussion is partly based on reconstructions and partly on runic material available to us today (as mentioned in Chapter 1, 1225 constitutes the onset of the use of the Latin alphabet in Sweden which supplanted the earlier *fuþark*). Some parts of the grammaticalization processes however elude us completely.

Interestingly enough, it is these earliest stages of grammaticalization that have received most attention from historical linguistics of all times and confessions. With very limited and partly undeciphered material the field for speculation is wide. Some speculations seem to be better founded and argued for than others, but unless we discover some lost documents they will necessarily remain speculations.

The chapter is organized as follows: in 3.2 the etymologies of the exponents of the category of definiteness are presented, including the weak adjectival inflection, the demonstratives *(h)inn* ‘that’ and *så* ‘this’, the numeral *en* ‘one’ and the existential determiner *någon* ‘some’. In 3.3 we take a closer look at the relative order of events, trying to establish the chronology of the grammaticalization processes. In particular, the time of cliticization of *(h)inn* is discussed. In 3.4 the occurrences of all exponents in the runic material from 800-1225 are presented.

3.2. Etymologies of the articles

In a study of a grammaticalization process it is important to establish which form undergoes grammaticalization, namely what its etymology is. In the case of the definite articles in Scandinavian languages, the task is neither simple nor uncontroversial.
The double, or more accurately, triple, definiteness in MS (see 1.5) mirrors several layers of linguistic innovation. The oldest of these is the formation of a new adjectival inflectional paradigm, the so-called weak inflection. This development, dated to the PGmc period, is common to all Germanic languages and has even been termed unique to them within the Indo-European language family (e.g. Ringe 2006); though it must be noted that similar tendencies were even present in Old Church Slavonic and in the Baltic languages.

The second layer of the morphologization of the category of definiteness is the formation of the postposed definite article, common to all Nordic languages, with the notable exception of the South Jutlandic dialects of Danish which have not developed a suffixed form of the definite article (e.g. Perridon 2002:1019). The suffixed article is most commonly derived form a distal demonstrative pronoun *hinn ‘that’, which is probably a cognate of the English ‘yon’.

The third layer of the multiple definiteness formation is the grammaticalization of the proximal demonstrative *sá (later *þæn) in the prepositional position. This process continues in OS, as witnessed by the oldest Swedish manuscripts and can therefore be tentatively dated at period I (1225-1375).

The indefinite article does not grammaticalize until the OS period. Its etymological source is the numeral *en ‘one’.

### 3.2.1. The weak adjectival inflection

The oldest grammaticalization of a definite exponent we know of today is the formation of a new adjectival inflectional paradigm, the so-called weak inflection, whereby the adjectives\(^{22}\) acquire a new set of endings, uniform with the n-stem nominal paradigm, as in (1):

\[
\begin{array}{lll}
\text{strong} & \text{weak} \\
\text{masculine} & *k\text{"ik\text{"a-}, fem.} & *k\text{"ik\text{"ō-} \\
\text{masculine} & *k\text{"ik\text{"a-an-, fem.} & *k\text{"ik\text{"ōn-} \\
\text{ }(Ringe 2006:169)
\end{array}
\]

This development, dated to the PGmc period, is common to all Germanic and several other Indo-European languages. In Old Church Slavonic there are the ‘long’ and ‘short’ adjectival paradigms, the long made up of an adjective and a cliticized demonstrative, where each element retains its original inflection in case and number.

---

\(^{22}\) Morphologically, adjectives are nouns, and in Indo-European have same declensions as the corresponding noun stems, thus *bonus* is inflected as *hortus*. Since they denote quality they usually appeared as attributes to nouns, e.g. *hortus magnus* (Prokosch 1939:259)
Remnants of this system can be found in modern Serbo-Croatian, and, though limited to nom and acc sg masc, in Slovenian as well (Krámský 1972:179-180, Lunt 2001). The weak declension implies the change of all adjective stems to n-stems—technically it consists of the addition of an n-determinant. This process has a counterpart in Latin. There the masculine adjectives could be changed to n-stems to denote permanent quality, and the new stems often formed proper names: *catus* ‘sly’, *Catō, -ōnis* ‘the sly one’ (Prokosch 1939:260). The term ‘individualizing’ is used (e.g. Krahe 1969, Ringe 2006) in the description of the weak inflection:


The nature of this individualization is thus to bring an object, entity or person amongst several similar objects, entities or persons into focus by naming the quality that sets it apart from the others. So for instance the ON *gamli maðr* ‘old-DEF man’ would correspond to ‘one of men, the old one’. This strategy disambiguates the referent for the hearer and makes it unique within some group of similar referents. The difference between the strong and the weak inflection can be illustrated by the following examples:

(3a) hans siukt ben
    his sick-ST leg
(3b) hans siuka ben
    his sick-W leg

(Delsing 1994)

In example (3a), the strong form of the adjective implies that it is used attributively; the meaning intended is probably ‘his leg, which, by the way, is sick’. In
the second example, the weak inflection has a contrastive function; by using the weak form the interpretation intended is more ‘the one of his legs that is sick’. The weak and strong paradigms are still used in this way in the classical Old Swedish texts. Consider the following example from YVL (ca 1280):

(4) Konæ  fi rígær  manni  fællir  hana  luct
Woman  kills  man  charge  her  closed-ST.NOM.FEM

hæræznæmpðin.
jury-DEF

(YVL FB:12)

‘If a woman kills a man, she shall be charged by a closed jury’.

The most probable meaning intended here is purely descriptive, not contrastive: ‘a jury that is closed’ not ‘this jury which is a closed one’ (out of a number of juries, some open). Thus, the strong form of the adjective is used, despite the fact that the NP is definite and overtly marked for definiteness by means of a definite article.

The combination of indefinite (strong) adjective + definite noun is seen by Berkov (1964) as a ‘new characteristic given to a known referent’, whereas that of definite (weak) adjective + definite noun is understood as a ‘known characteristic of the known referent’. Such examples are still found in Modern Icelandic, where the following combinations are available:

(5a) góður  maður
good-ST  man

(5b) góði  maður
good-W  man

(5c) góði  maðurinn
good-W  man-DEF

(5d) góður  maðurinn
good-ST  man-DEF

(based on Naert 1969)

The choice of strong vs weak form of an adjective is meaningful, as illustrated by the different interpretations of the adjective in (6a) and (6b).

(6a) Ég  horfði  upp  í  bláan  himininn.
I  looked up into blue-ST  sky-DEF

‘I looked at the sky that was blue’.
In (6a) ‘the sky’ is definite since it is a unique referent. Its property of being blue is however not inherent—the sky may be e.g. grey or black. Therefore the strong form of the adjective blá ‘blue’ suggests that the colour of the sky is new information. In example (6b) both the car and its colour are elements the hearer is assumed to be familiar with perhaps because the blue car in question has been discussed before. Such use of the weak vs strong adjectival inflection certainly bears some parallels to the restrictive vs non-restrictive subordinate clauses.

(7a) I looked at the sky, which was blue (=one sky, its colour is additional information).

(7b) I looked at the car which was blue (=more than one car, the one I looked at was the blue one).

The use of the weak form implies that there is a choice of more than one potential referent and the one described by the quality is to be chosen (‘individual-izing’ function). In Icelandic, it is regular to the point of ungrammaticality of the following:

(8) Rautt/??rauða nefið á honum glóði í myrkrinu.
red-ST/??W nose-DEF on him glowed in dark-DEF
‘His red nose glowed in the dark’.

The ungrammaticality of the weak form of the adjective in (8) lies in the fact that it would imply that the person had more than one nose and it was the red one that glowed in the dark or at least that there is a contrast between the red nose and some other nose.

In Swedish, the weak adjectival inflection gradually turns into an agreement phenomenon and becomes obligatory in any definite NP, independent of what the other exponents of definiteness are. Therefore there is no difference similar to the difference made in Icelandic:

(9a) Jag tittade upp i den blåa himlen.
‘I looked up to the blue sky’.
Weak adjectival inflections develop in several Indo-European languages. The Germanic innovation, for all adjectives excepting a few indeclinable, is that there are two paradigms available: the weak and the (original) strong. The weak is identified with definite meaning though the number of contexts for it are limited. The modern usage is purely a concord phenomenon, guided by syntactic considerations rather than semantic ones.

### 3.2.2. Demonstratives

Typologically it is usual for definite articles to be derived from demonstrative pronouns and indefinite articles from the numeral ‘one’ (see chapter 2.5.1 and 2.5.2). So is the case in Swedish. There were three, perhaps even four, demonstrative pronouns in the Old Nordic pronominal system (Noreen 1913, Krahe 1948). Those included: *sá* (nominative masculine, later substituted by the accusative form *þen*), *hinn* (or *inn*), and the compounded *sjá* (later *penna*). The forms of the pronouns are given in the tables 5a, 5b and 5c respectively (*inn* has the same inflectional forms as *hinn*).

**Table 5a. The inflectional paradigm of (h)inn**

<table>
<thead>
<tr>
<th>Number</th>
<th>Case</th>
<th>M</th>
<th>F</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>NOM</td>
<td><em>hinn</em></td>
<td><em>hin</em></td>
<td><em>hit</em></td>
</tr>
<tr>
<td></td>
<td>GEN</td>
<td><em>hins</em></td>
<td><em>hinnar</em></td>
<td><em>hins</em></td>
</tr>
<tr>
<td></td>
<td>DAT</td>
<td><em>hinum</em></td>
<td><em>hinni</em></td>
<td><em>hinu</em></td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td><em>hinn</em></td>
<td><em>hina</em></td>
<td><em>hit</em></td>
</tr>
<tr>
<td>PL</td>
<td>NOM</td>
<td><em>hinir</em></td>
<td><em>hinar</em></td>
<td><em>hin</em></td>
</tr>
<tr>
<td></td>
<td>GEN</td>
<td><em>hinna</em></td>
<td><em>hinna</em></td>
<td><em>hinna</em></td>
</tr>
<tr>
<td></td>
<td>DAT</td>
<td><em>hinum</em></td>
<td><em>hinum</em></td>
<td><em>hinum</em></td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td><em>hina</em></td>
<td><em>hinar</em></td>
<td><em>hin</em></td>
</tr>
</tbody>
</table>

The demonstrative pronoun *(h)inn* is found in the eastern languages (Swedish and Danish) almost exclusively as *hinn* (there are five runic inscriptions where the form *inn* appears, U 226, U 1146, Hs 21, Sö 41, Sö 125, see *Rundatabasen*), whereas in Icelandic two forms are in variation: *hinn* and *inn/enn*. There they seem to be interchangeable in all contexts, though a more detailed study is called for.
This variation in texts has not been given sufficient prominence in the literature, though there have been many attempts to establish the etymologies of both *hinn* and *inn*. Some explanations for the existence of two such similar pronouns are morphological and some phonological. According to the former, *hinn* is the result of a composition of two pronominal demonstrative stems: *inn* and *h*- , both of which can also be found in the 3rd person personal pronouns (sg) *han, hon*. This demonstrative survives in modern German *heute* (lit. ‘this day’, Krahe 1948:67). Such a view of the origin of *hinn* is held among others by Krahe (1948:67), No- reen (1913:180 and 190), de Boor (1922:173 and 178), Syrett (2002:721). A minority of authors give a phonological or phonetic explanation to the presence of *h* in *hinn* (Gjerdman 1924:137, Jónsson 1901:315, Neckel 1924:407-412). For a detailed account of the disagreement see Perridon 1989:129-135.

Table 5b. The inflectional paradigm of *sá*

<table>
<thead>
<tr>
<th>Number</th>
<th>Case</th>
<th>M</th>
<th>F</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NOM</td>
<td><em>sá</em></td>
<td><em>sú</em></td>
<td><em>þat</em></td>
</tr>
<tr>
<td></td>
<td>GEN</td>
<td><em>þess</em></td>
<td><em>þeir(r)</em></td>
<td><em>þess</em></td>
</tr>
<tr>
<td></td>
<td>DAT</td>
<td><em>þeim</em></td>
<td><em>þeir(r)i</em></td>
<td><em>því</em></td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td><em>þann</em></td>
<td><em>þá</em></td>
<td><em>þat</em></td>
</tr>
<tr>
<td>SG</td>
<td>NOM</td>
<td><em>þeir</em></td>
<td><em>þær</em></td>
<td><em>þau</em></td>
</tr>
<tr>
<td></td>
<td>GEN</td>
<td><em>þeir(r)a</em></td>
<td><em>þeir(r)a</em></td>
<td><em>þeir(r)a</em></td>
</tr>
<tr>
<td></td>
<td>DAT</td>
<td><em>þeim</em></td>
<td><em>þeim</em></td>
<td><em>þeim</em></td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td><em>þá</em></td>
<td><em>þær</em></td>
<td><em>þau</em></td>
</tr>
</tbody>
</table>

* later superceded by the acc forms *þan, þa, þat*
The nominative forms given in table 5b occur mainly in runic inscriptions and, sporadically, in the oldest Swedish legal texts. As they bear no resemblance to other forms of the paradigm they are most probably a remnant of yet another pronominal paradigm, which is found in other Indo-European languages (e.g. Slavic s-forms; Polish do siego roku). This paradigm is highly weakened in other Indo-European languages as well. Originally the s-pronouns determined an entity found ‘close to the hearer’ and contrasted with d-pronouns (close to the speaker) and x-pronouns (far from both the speaker and the hearer). This tripartite system is preserved in some Indo-European languages, e.g. Armenian (see Chapter 2), Ibero-Romance, some Italian dialects in Southern Italy. Whether they share a common ancestry is unclear (Bauer 2007:105, cf. Brugmann 1911:310-47, esp. 346-7).

Table 5c. The inflectional paradigm of sjá (þenna)

<table>
<thead>
<tr>
<th>Number</th>
<th>Case</th>
<th>M</th>
<th>F</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>NOM</td>
<td>sjá*</td>
<td>sjá*</td>
<td>þetta</td>
</tr>
<tr>
<td></td>
<td>GEN</td>
<td>þessa</td>
<td>þessa</td>
<td>þessa</td>
</tr>
<tr>
<td></td>
<td>DAT</td>
<td>þeima</td>
<td>þessar(r)i</td>
<td>þessa</td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td>þenna</td>
<td>þessa</td>
<td>þetta</td>
</tr>
<tr>
<td>PL</td>
<td>NOM</td>
<td>þessir</td>
<td>þessar</td>
<td>þessi</td>
</tr>
<tr>
<td></td>
<td>GEN</td>
<td>þessa</td>
<td>þessa</td>
<td>þessa</td>
</tr>
<tr>
<td></td>
<td>DAT</td>
<td>þessum</td>
<td>þessum</td>
<td>þessum</td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td>þessa</td>
<td>þessar</td>
<td>þessi</td>
</tr>
</tbody>
</table>

*later superceded by the acc forms þún, þu

The difference between the pronouns (h)inn and sá had most probably to do with marking of distance: (h)inn corresponded to English ‘yon’ (distal deixis), while sá marked the proximal deixis. In its paradigm remnants of a yet another proximal demonstrative are found. Sjá functioned as a reinforced deixis (a compound of demonstrative sá and intensifying -si).

The tripartite system which we can discern in the (h)inn and sá/pen paradigms can be illustrated by the Homeric Greek and Latin demonstratives (after Bauer 2007:105):

(11a) ó ‘this’ (‘this, close to me’, 1st person)
(11b) οὗτος ‘that’ (‘that, close to you’, 2nd person)
(11c) ἐκείνος ‘that’ (‘that, close to him’, 3rd person)
When ὁ grammaticalized into a definite article, a demonstrative particle was added to the original pronoun rendering ὅδε and re-establishing the original trichotomy (Bauer 2007:106).

The Latin demonstrative system was a tripartite as well, as illustrated in the following:

(12a) hic ‘this’ (1st pers., EGO)
(12b) iste ‘that’ (2nd pers., TU)
(12c) ille ‘that’ (3rd pers., ILLAE)

Interestingly, all of the demonstratives are compounds. Additionally, Latin had an anaphorically used element is, referring to an entity just mentioned or about to be mentioned but without a deictic value (Bauer 2007:113). This is also the element which disappears from the language, remaining only in some fixed expressions.

### 3.2.3. Numerical ‘one’

The indefinite article in Swedish is derived from the numeral enn ‘one’, PGmc *aina- from IE. *oi-no (Lat. unus, Greek dial. oinós). Its grammaticalization does not start until the beginning of the OS period and its occurrences in the runic material are exclusively as a numeral (see 3.4)

<table>
<thead>
<tr>
<th>Number</th>
<th>Case</th>
<th>M</th>
<th>F</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>NOM</td>
<td>enn</td>
<td>en</td>
<td>ett</td>
</tr>
<tr>
<td></td>
<td>GEN</td>
<td>ens</td>
<td>enna</td>
<td>ens</td>
</tr>
<tr>
<td></td>
<td>DAT</td>
<td>enom</td>
<td>enne</td>
<td>enom</td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td>enn</td>
<td>ena</td>
<td>ett</td>
</tr>
</tbody>
</table>

There are two other elements linked with the grammaticalization of the indefinite article. The first is the någon ‘some’, used in nonspecific indefinite contexts (see Chapter 1.5). Its etymology clearly indicates that it is a nonspecific marker; någon is a relatively late formation and unlike (h)inn, så, en or sjá does not trace back to a single IE element. It was formed in the AN/ON time (Hellquist 1957:713) as a contraction of the phrase ne wait ek hwariR, literally ‘not know I which’. In OS sources it appears as nokorr.

Another grammaticalized formation connected with the grammaticalization of the indefinite article is MS ingen ‘none’, regularly used in negated utterances (without the negation occurring a second time), e.g.
(13) Jag köpte ingen bil.
I bought none car.
‘I didn’t buy a car’.

The form is a coalescence of *en* (the numeral and later indefinite article) and a particle -ghi(n), which originally had a generalizing meaning (Hellquist 1957) and later a negating one
23. In subordinate clauses it is split into the negation marker *inte* and *någon*.

(14) Jag vet att han inte har köpt någon bil.
I know that he not has bought NGN car
‘I know that he hasn’t bought any car’.

### 3.3. The chronology of the grammaticalizations

As we stated in 3.2, the MS definite NP comprises three different grammaticalizations—of weak adjectival inflection, demonstrative pronoun *(h)inn* and demonstrative pronoun *sá* (*pen*). These grammaticalizations are not simultaneous in the historical development—the formation of the weak adjectival inflection belongs to the PGmc period (see 3.2.1), while the demonstratives grammaticalize later.

Of all ingredients of the grammaticalization of (in)definiteness in the Scandinavian languages, including Swedish, that of *(h)inn* has attracted most scholarly attention. This interest is natural, since one would wish to explain why the definite NPs in Scandinavian take the form of e.g. *bok-en* ‘book-DEF’ rather than *den bok* ‘the book’, as they do in closely related and neighbouring German, Frisian and Dutch dialects
24, as well as in English. The studies of the developments are based on the West Nordic sources (see Chapter 1.2), since these are some 1000 years older than the East Nordic (Swedish and Danish) ones.

The Swedish definite article is a suffix. It is unusual and within the IE language family only found in Bulgarian, Romanian and Albanian. Its formation followed most likely followed the cline of grammaticalization:

(15) lexeme > clitic > affix

Cliticization of the demonstrative predates the Swedish sources (from 1225 onward) and may belong to the common history of the Scandinavian languages. The form found in OS already shows affix-like properties (for a discussion of the

---

23 Compare also MS *aldrig* ‘never’—a coalescence of *alder* ‘time’ and the same particle -ghi(n), literally ‘not in any time’.

24 And, as we have mentioned before, in some Jutlandic dialects of Danish.
The chronology of the grammaticalizations

status of -IN as either clitic or affix in ON, which is valid even for OS, see Faarlund 2007, Abraham and Leiss 2007 and a critique of both in Börjars and Harris 2008).

The precondition for the definite suffix/enclitic to arise is the postposition of the demonstrative—as the position of inflectional material mirrors the original constituent order within the NP. Only when placed after the noun can the demonstrative be enclitically attached to it.

Table 7. The definite nouns in Old Swedish

<table>
<thead>
<tr>
<th>Number</th>
<th>Case</th>
<th>m ‘day’</th>
<th>f ‘journey’</th>
<th>n ‘ship’</th>
</tr>
</thead>
<tbody>
<tr>
<td>sg</td>
<td>nom</td>
<td>daghr-in</td>
<td>færþ-in</td>
<td>skip-it</td>
</tr>
<tr>
<td></td>
<td>gen</td>
<td>daghs-ins</td>
<td>færþ-inna(r)</td>
<td>skips-ins</td>
</tr>
<tr>
<td></td>
<td>dat</td>
<td>daghi-ins</td>
<td>færþ-inne</td>
<td>skipi-nu</td>
</tr>
<tr>
<td></td>
<td>acc</td>
<td>dagh-in</td>
<td>færþ-ena</td>
<td>skip-it</td>
</tr>
<tr>
<td>pl</td>
<td>nom</td>
<td>dagha-ni(r)</td>
<td>færþe-na(r)</td>
<td>skip-in</td>
</tr>
<tr>
<td></td>
<td>gen</td>
<td>dagha-nna</td>
<td>færþa-nna</td>
<td>skipa-nna</td>
</tr>
<tr>
<td></td>
<td>dat</td>
<td>daghum-in</td>
<td>færþom-in</td>
<td>skipum-in</td>
</tr>
<tr>
<td></td>
<td>acc</td>
<td>dagha-na</td>
<td>færþe-na(r)</td>
<td>skip-in</td>
</tr>
</tbody>
</table>

Two major hypotheses have been put forward regarding the cliticization.

(16a) \( N(h)inn \text{ Adj} \rightarrow N-\text{inn} \text{ Adj} \) (Grimm 1837, Delbrück 1916)

(16b) \( N(h)inn \rightarrow N-\text{inn} \) (Nygaard 1905, Pollack 1912)

Hypothesis (a) assumes that the demonstrative following the noun belonged to an adjective phrase with the adjective in the weak form, as in (17).

Hypothesis (b) is the more general of the two; it assumes that a demonstrative could be postposed and in this position unstressed, which leads to the gradual loss of deictic force and syntactic independence.

(17) maðr inn gamli
    man.NOM that old.W

25 ‘Today’s morphology is yesterday’s syntax’ (Givón 1971:413).
26 A third reconstruction assumes the postposition of \((h)inn\) analogical to MS emphatic formation such as bonden, han var rik ‘the farmer, he was rich’ (Gjerdman 1924). It is problematic since the suffixed article does not show forms of the personal pronoun paradigm. The reconstruction has largely been ignored in later literature, apart from Perridon 1989:143-149, who finds it adequate and explanatory of the origins of articles by subjects.
There is little doubt that the weak adjectives were often reinforced by a demonstrative and that it was dominantly *(h)inn* to begin with. In this function, *(h)inn* has been termed an adjective-article (e.g. Stroh-Wollin 2009).

Reconstruction (a) is different from the more general (b), since it assumes a reanalysis, in fact a re-bracketing, process. According to (a), the original structure of the NP was:

(18)  

+[maðr] [inn gamli]  

The demonstrative and the weak adjective formed one unit, which could be placed either before (19a) or after (19b) the noun, and could also be detached from it by other elements (19c)27.

(19a)  

and now could DET old-W wolf-DEF come at  
gullinu  
gold-DEF  
‘And now the old wolf managed to get to the gold’ (Musinowicz 1911:60)28  

(19b)  

father Haldan-GEN DET-GEN mild-W and DET-GEN  
matar illa  
food bad-GEN  
‘Father of Halfdan the mild and stingy of food’ (Musinowicz 1911:57)  

(19c)  

Isleif-GEN Gizor-GEN son-NOM DET-GEN  
huita  
white-W was ordained to bishop-GEN  
‘Isleif, son of Gizor the white, was ordained bishop’ (Musinowicz 1911:57)  

However, due to a tendency to postpose attributes and modifiers, the adjective phrase was more often found after the noun. The structure in (18) came with time to be reanalyzed as (20).

(20)  

+[maðr inn] gamli  

(Delbrück 1916)

27 All following examples are from ON texts.

28 Examples quoted after Musinowicz (1911) retain the spelling given there, which is not normalized.
A number of objections to hypothesis (a) have been raised, mainly by Pol-
lack (1912) and before him by Nyggard (1905), who proposed an the alternative
reconstruction (b), in which the postposition of \((h)inn\) was assumed as well, but
not a \((h)inn\) that belongs together with a weak form of an adjective. Their major
argument against hypothesis (a) is that the postposed adjective phrases usually
accompanied proper names, while the appelatives were modified by the preposed
phrase. The argument is largely based on Musinowicz 1911, who in his study of
the position of the adjective attribute found only 4 instances of a postposed weak
adjective with \((h)inn\) where the noun was not a proper name out of 102 examples
(Musinowicz 1911:56).

Though Musinowicz’ study is thorough and impressive, its results are hardly
conclusive for the reconstruction. Firstly, the fact that, in his material, \((h)inn + \)
adjective are postposed almost only after proper names does not exclude tendencies
to cliticization. Such tendencies are quite clear in examples such as (20), where
the demonstrative seems to be understood by the writer as part of the noun rather
than a separate word\(^{29}\).

(20)  Erikinum    hælghæ
     Erik-DEF holy-W
     (Kock 1919:98-99, also quoted in Börjars and Harris 2008:297)

Secondly, there is a general agreement that the tendency to postposition of
nominal modifiers was stronger in the older stages of the development of the
Scandinavian languages, and that it gradually wears off. The study of ON is based
on material where both pre- and postmodification is possible, with a growing ten-
dency to premodification. It was a time where the old preference for postmodifi-
cation was being abandoned in favour of premodification. The cliticization of
\((h)inn\) is of an earlier date than the texts studied, which is clear from exam-
pies such as (21a) and (21b), where it occurs or in conjunction with the pre- or
postposed \((h)inn\).

(21a)  En    síðan leiddi hann Flóres út i pann    enn
    but later led he Flores out in DET DET

    góða    eplagarðinn
    good-W applegarden-DEF
    ‘But later he led Flores out to the pleasant apple orchard’.

\(^{29}\) Naturally, we may only assume that when the OS or ON writer wrote two elements as one
word they formed a unit in his grammar. However, we sooner see the opposite tendencies: to write
separately words that in fact are a compound (they are regularly adjacent and no intervening ele-
ments come between them, e.g. \(kirkio gaardh\) ‘churchyard’). So, if the writer was more prone to
write compounds as if they were separate words, when he does write words together they most
probably function as one for him.
(21b)  sem þau sváfu í hjónin hin
which they slept in married couple-DEF DET

ungu
young-W
‘which they slept in, the young married couple’
(both after Musinowicz 1911:60)

Finally, it should be noted that the reconstruction by Nygaard and Pollack itself suffers from lack of data. The demonstrative (h)inn is found in an adnominal position only once in the runic material before 800, on the whet-stone from Strøm in Norway, dated at ca 600 AD.

(22) wate hali hino horna
wet stone DET horn
‘Let the horn wet this stone!’

Otherwise, only sá or sjá is found in this position in the runic inscriptions, either preceding or following the noun, typically stein ‘stone’ on which the inscription is carved.

(23a) þat azina (By stone, ca 500 AD, Norway)
‘this stone slab’

(23b) stein saR (U 10)
‘stone this’

(23b) þoriR ok hroða letu ræisa stæin þennsa (U 429)
‘Thore and Hroda had this stone raised’

(23c) astrið let ræisa þenna stæin (U 238)
‘Astrid had this stone raised’

This is understandable when we consider that the proximal deictic demonstrative is more likely to appear in inscriptions meant to read e.g. ‘this stone’. However, definite articles could also be used deictically in OS, as in (24) or in later runic inscriptions (25).

(24) som her star i bokenne scriuit
which here stands in book-DEF written

‘Which is written here in the/this book’.
(after Larm 1936:160 and Perridon 1989:133)
The runic material as well as the earliest ON texts indicate \((h)inn\), only sporadically occurring by nouns, as the adjective article and \(sá\) or \(sjá\) as the adnominal demonstratives.

We now turn to the question why weak adjectives should be accompanied by demonstratives at all, as they seem to be in many contexts, both in pre- and postposition. Hodler (1954) argues that the demonstrative was used in this position to reinforce the weakened definite reference of the adjective. However, the weak form could be used on its own, particularly in preposition. Another striking fact about the distribution of the postposed attributes is that they could easily be separated from the head (the noun) by intervening elements. However, when separated, the weak adjective required a demonstrative pronoun (a strong adjective could not be postposed in the same way), as in (26a) and (26b).

\[(26a)\]  
\[
\begin{array}{llllllllll}
\text{ok} & \text{má} & \text{hann} & \text{þá} & \text{ráþa} & \text{gullinu} & \text{því} & \text{inu} \\
\text{and} & \text{may} & \text{he} & \text{then} & \text{rule} & \text{gold-DEF-DAT} & \text{DET} & \text{DET} \\
\end{array}
\]

\text{mikla}  
\text{large-W}  
‘And he may then have this large amount of gold at his disposal.’

\[(26b)\]  
\[
\begin{array}{llllllllll}
\text{þetta} & \text{hit} & \text{sama} & \text{setr} & \text{hann} & \text{í} & \text{miðju} \\
\text{DET} & \text{DET} & \text{same} & \text{place} & \text{he} & \text{in} & \text{middle-W} \\
\end{array}
\]

\[
\begin{array}{llllllllll}
\text{því} & \text{landjaldinu} & \text{hinu} & \text{stóra} \\
\text{DET} & \text{tent-DEF-DAT} & \text{DET} & \text{big-W} \\
\end{array}
\]

‘There in the middle he placed the tent, the big one’  
(Musinowicz 1911:60)

An alternative definition of the function of the demonstrative would be of an element linking the NP together. A recent study of the origin of the definite article in Swedish, Stroh-Wollin 2009, defines its function along these lines. In Stroh-Wollin’s terminology the pre-adjectival \((h)inn\) has a mediating function (Stroh-Wollin 2009:8f.) whose role is difficult to explain. There are parallels from other languages in recent research, described in Himmelmann (1997) and Diessel (1999), where a similar phenomenon is identified as a ‘linking article’. Its role is easier to understand when we realize that the NP could be split and the postposed material separated from the main noun by other elements (compare the Modern Swedish: \textit{Handeln ökar med Frankrike} ‘Trade is rising with France’ instead of the
equally grammatical *Handeln med Frankrike ökar* ‘Trade with France is rising’; in the first example the NP is divided by the verb). In such cases the demonstrative would serve as a linking device, bringing the phrase together. As the postposed linker became obligatory, even when no intervening elements occurred within the NP, it appeared directly after the noun with increasing frequency, allowing the cliticization process to commence.

In its function of adjective article, *(h)inn* may be exchanged for *sá*, with no apparent difference in meaning, as illustrated by the following example:

(27) (..) resþi sten (..) at Óþinkor (..) þan dyra ok hin

*drottinfasta*

lord-loyal

‘Raised the stone in memory of Óþinkor, valued and loyal to his lord.’

However, the presence of both *ben* and *hin* in the same context within one inscription may suggest some contrast between the phrases possibly referring independently to two of the people mentioned (Johnsen 1976:58). The meaning of *hinn* would then incorporate ‘the foregoing’ or ‘the other’, and it is in fact found with this meaning in the oldest texts (the adversative *hinn*). However, the linguistic material is too scanty to render any conclusive evidence of either interpretation.

The regional distribution of *(h)inn* and *þen* in this position is not identical; while *(h)inn* dominates in the western varieties (Icelandic and Norwegian), *þen* is more common in the eastern ones (Danish and Swedish). Also, the variation in form between *inn* and *hinn*, known from the Icelandic manuscripts, is not attested in Sweden, where only *hinn* is to be found30. The variation in Icelandic manuscripts has to the author’s knowledge not been studied systematically and is generally assumed to be free (e.g. Sprenger 1977).

All authors discussing the subject agree at least on one point: that the adjective article is older than the clitic -IN. They also seem to agree that the adjective article is a result of weakening of the definite meaning of the weak adjectival inflection. From that point on, opinions differ. For many, like Delbrück, Nygaard and, in recent years Ringe, the weak adjectival inflection is similar in function to the definite article. For Hodler, however, it is pure nominalization, and as such possible to use with definite, indefinite or generic meaning31. For Delbrück the ad-

---

30 Apart from the five inscriptions mentioned earlier.

31 Diese Ansicht ist—kurz gesagt—falsch. Das schwache Adjektiv ist .. eine Substantivierungsform ... Wie jedes Substantiv konnte eine solche Bildung bestimmte, unbestimmte und generelle Bedeutung haben (Hodler 1954:76). Hodler’s point-blank rejection of the definite meaning of the weak adjectival inflection is criticized among others by Møller, who writes that ‘det er dog vel rimeligt, at der har ligget noget ‘bestemt’ i den svage form, som fra at være et individualiserende, personkarakteriserende substantiv blev opfattet som adjektiv’ (Møller 1956:121)
jective article and the clitic are the same, given a time difference: his reconstruction is based on the postposition of an adjective phrase with adjective article. For others the two processes are separate and the postposition of (h)inn has nothing to do with its contemporary function as an adjective article (Pollack, Nygaard).

It must be noted that the adjective article in ON is definitely not anaphoric but must serve some other function. When an anaphoric reference is made another demonstrative is added.

(28a) þann inn alsvinna iotvn
DETT DETT wise-W giant

(28b) sa inn fráni ormr
DETT DETT gleaming-W snake
(examples after Nygaard 1905:87)

Delbrück’s reconstruction builds upon postposition of an adjective phrase, composed of a demonstrative (h)inn and the weak form of an adjective whereby the demonstrative gradually cliticizes onto the preceding noun and is reinterpreted as belonging together with the noun rather than the adjective. The reconstruction was criticized heavily almost as soon as it was put forward, among others by Nygaard (1905) and Pollack (1912), who instead assumed postposition of the demonstrative alone. A study of the phrases consisting of N followed by (h)inn and a weak form of an adjective revealed that they consist most frequently of a proper name and not a common noun32 (Musinowicz 1911). This has posed a serious challenge to the Delbrück hypothesis.

Stroh-Wollin (2009) assumes that the pre-adjectival (h)inn and the suffix -IN are instances of two separate grammaticalization processes, although from a common source (see also Faarlund 2007). In the pre-adjectival position (h)inn competes with þen33. Seeing pre-adjectival (h)inn and the suffix -IN as two separate grammaticalizations enables Stroh-Wollin to entirely escape the discussion on

32 The adjective phrase after the proper name could have either a restrictive or a descriptive meaning. This is illustrated by examples from runic inscriptions: Balli hinn rauði ‘Balli the red’ (Vs 15), possibly as opposed to other carvers by name of Balli and Kristr hinn hölgi ‘the holy Christ’ (U 391) (Stroh-Wollin 2009:8).

33 Stroh-Wollin uses the term ‘article’ for all three forms assuming thereby: 1. that there is more than one definite article in the language (interestingly enough she hesitates to term -IN an article); 2. that all forms are fully grammaticalized articles, which is a controversial viewpoint. In the present study no such assumptions are made and when relating Stroh-Wollin the term ‘demonstrative’ is used, which is not her intention. One of the points she makes is that the suffix is not a grammaticalization of a free article (Stroh-Wollin 2009:5), but even though I disagree with the term ‘article’ in regard to all OS elements I think her arguments still hold if (h)inn is seen as just a pre-adjectival demonstrative. To term it an article puts an unnecessarishly sharp border in the cline of grammaticalization.
Delbrück. In her view cliticization of \((h)inn\) is earlier than its grammaticalization in the pre-adjectival position (Stroh-Wollin 2009:7) and therefore construction of the type \(maðr\inn\,goði\), which is the cornerstone of the Delbrück analysis, cannot possibly be its source.

Another possible explanation is that the weak adjectival inflection had already lost much of its contrastive meaning, which now had to be reinforced by the demonstrative.

The date of the cliticization of \((h)inn\) still eludes us. Several different hypotheses have been put forward as to why the article is suffixed. Also, different authors have proposed different dates for article formation. These are presented in table 8.

<table>
<thead>
<tr>
<th>Proposed dating</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN/ON</td>
<td>Delbrück 1916(^{34}), Neckel 1924</td>
</tr>
<tr>
<td>500-1100</td>
<td>Syrett 2002</td>
</tr>
<tr>
<td>ON</td>
<td>Gjerdman 1924, Larm 1936,</td>
</tr>
<tr>
<td></td>
<td>Barnes (cf. Börjars and Harris 2008:295f)</td>
</tr>
<tr>
<td>The Viking Age (just before the legal texts written down)</td>
<td>De Boor (and von Friesen)</td>
</tr>
<tr>
<td>1000-1150</td>
<td>Seip 1958(^{35})</td>
</tr>
<tr>
<td>With the first written texts</td>
<td>Braunmüller</td>
</tr>
<tr>
<td>Ca 1200</td>
<td>Hansen 1927(^{36})</td>
</tr>
<tr>
<td></td>
<td>Delsing 2002(^{37})</td>
</tr>
</tbody>
</table>

\(^{34}\) Ich bin nicht mehr der Meinung, daß der Schlußartikel in der Sondergeschichte des Altisländischen entstanden sei, sondern glaube seinen Ursprung in die urnordische Zeit verlegen zu sollen. (Delbrück 1916:4)

\(^{35}\) Vi har ingen sikkerhet for at etterhengt artikkel er urnordisk. At den ikke fins i runeinskrifter før i en svensk ca. 1040, i noen norske fra 1100 av, taler nærmest mot det. .. Den gamle oppfatning at den etterhengte artikkel er utviklet i vikingtiden, .. har gode grunner for seg. (Seip 1958:240).

\(^{36}\) Man maa huske paa, at talesproget ikke havde nogen vidare brug for artiklen, saa længe man havde forbindelsen af demonstrativ og subst. Det er skriftsproget, der her—som i andre sprog—fremmer udviklingen af en artikel for trykket. Først da de nordiske sprog faar sig en litteratur, og især da der begynder at danne sig en skriftsprogsform, begynder artiklen at brede sig og blive alminderligere brugt. (Hansen 1927:170)

\(^{37}\) I am inclined to believe that the article developed as an innovation in the 13th century. (Delsing 2002:939)
3.4. The runic period (ca 800-1225)

As mentioned before, in WN languages both *inn* and *hinn* appear as free morphs, though no clear difference in meaning can be found between the two (Delbrück 1916). In EN languages on the other hand, there are no instances of *inn* and *hinn* is the only free form. If we then take *hinn* to be of a later date than the earliest instances of the definite article we must conclude that by the time Swedish became an independent language the article had already been formed.

Such an early date proposed by some scholars (Delbrück 1916, Neckel 1924, Barnes (cf. Börjars and Harris 2008)), is seriously challenged by the fact that the earliest instances of the definite suffix are three debatable examples form the runic material and the twenty odd examples from *Äldre Västgötalagen* (see Chapter 4.3). One of the runic examples is a Danish inscription from Tornby, dated at ca 1200 and attributed to a Norwegian rune master (*Rundatabasen*).

\[
\text{(29) } \text{um morhen(e)n on morning-DEF} \quad \text{(DR 169)} \\
\text{‘(..) on the morning’}
\]

The other two are the Swedish inscriptions U 644 and U 669, where the form *antini* ‘the spirit’ can be found.

\[
\text{(30) } \text{kup heabi ontini God help soul-DEF} \quad \text{(U 644)} \\
\text{‘May God help the soul’}.
\]

The problem with U 644 is that the reading is uncertain (Delsing 2002:939). U 669 is lost altogether so we need rely on the drawings of it.

Postposition of the demonstrative was necessary for it to cliticize as an enclitic. The earliest sources show that NPs started with a noun and all attributes were postposed, as illustrated by a recurrent formula on runic inscriptions:

\[
\text{(31) } \text{NN raesti stein þenna aftir NN faður sin goðan} \\
\text{NN raised stone this after NN father his good} \quad \text{(Stroh-Wollin 2009:16)}
\]

From 800 AD the language continuum in Scandinavia begins to split into two distinct varieties: the West Nordic (whose descendants are Norwegian, Icelandic and Faroese) and East Nordic (whose descendants are Danish and Swedish).

38 The lost stone was a complement to U 668, which was raised by two men to honour their lost brother. The end of the inscriptions reads: *kup hialbi salu* ‘may God help soul’, with soul unmarked. It is strange that the lost stone (carved in memory of the father) should read *antini* ‘the spirit’.
The regional distribution of (h)inn and þen in this position is not identical; while (h)inn dominates in the western varieties (Icelandic and Norwegian), þen is more common in the eastern ones (Danish and Swedish). Also, the variation in form between inn and hinn, known from the Icelandic manuscripts, is not attested in Sweden, where only hinn is to be found. The variation in Icelandic manuscripts has to the author’s knowledge not been studied systematically and is generally assumed to be free.

One of the earliest examples of a Swedish runic inscription is the Rök stone (Ög 136), dated at ca 800. The inscription is relatively long (the stone is carved on all sides) and only partly read. There is still some amount of disagreement as to the correct interpretation of the text and a new reading has been recently proposed (Ralph 2007).

The inscription starts with the phrase aft umuþ stonta runaR þaR ‘after Vemund stand runes these’, with the postposed demonstrative sá. A bit further on in the text we find the first occurrence of hinn from the Swedish territory, in the phrase raipiaurikR hin þurmuþi ‘rode Tjodrik the brave’.

The indefinite article develops in Swedish in the 13th and 14th century; en ‘one’ is therefore not found in runic inscriptions in functions other than the numeral and there is only one instance from Sweden, G 136 †$, dated at 1000-1150 AD (Rundatabasen).

One instance of en could be interpreted as the specific marker rather than the numeral, namely U 102.

The numeral interpretation is proposed by both Bugge and Kock (see Leijström 1934:150), though on different grounds, Bugge sees a contrast between one bridge and two sons, while Kock sees a weak form of en (compare MS ena), which is to be related to ikidhora to emphasise that the bridge was raised on individual initiative (in contrast to bridges raised by the whole village, a custom

---

39 Apart from U 226, U 1146, Hs 21, Sö 41, Sö 125.
familiar from Uppland). Leijström, however, interprets "aina" in U 102 as devoid of numeral meaning and signifying indefinite meaning (*indefinit betydelse synes foreligga*, Leijström 1934:150ff.)

### 3.5. Conclusions

The process of definiteness grammaticalization in Swedish before 1225 may thus be sketched in the following way:

- **Proto-Germanic:** formation of the weak adjectival inflection
- **Proto-Nordic/Ancient Nordic:** formation of the ‘adjective article’—a growing tendency to mark weakly inflected adjectives additionally with a demonstrative
- **Old Nordic:** cliticization of the demonstrative to the noun / competition between *hinn* and *sá* as adjective article

The research so far has mainly concentrated on the earliest stages of grammaticalization of the definite article: its origin (in *inn* or *hinn*), the relationship between *hinn* and *inn*, means of development (postposition of a demonstrative or an adjective phrase) and the age of the development. No final answers have been given to any of these questions. In the following chapter we will study the development in the OS texts.
4 Grammaticalization of the definite article

4.1. Introduction

We have seen in Chapter 3 that scholarly interest in the formation of definiteness in Swedish (and other Scandinavian languages) focused on the very first stages of grammaticalization of the postposed definite article. There has been little interest in the development of the form beyond its cliticization and yet it can hardly be claimed that the definite article is fully developed in OS religious and profane prose\(^{40}\). Admittedly, the clitic -IN is already found in the oldest text, AVL, and we have no traces of the cliticization process. It is however largely missing in unique and generic contexts and is, more interestingly, still not fully obligatory in the textual functions either.

In this chapter we will look at its gradual rise in frequency and its spread through the textual functions, direct and indirect anaphora, and beyond them, to the non-textual ones—its use with unique and generic referents. We will also look closely at uses of -IN that seem anomalous from the MS perspective.

4.2. General results

General results of the frequency search for -IN in OS texts are collected in Table 9. They are also compared with the results for a MS narrative.

<table>
<thead>
<tr>
<th>-IN</th>
<th>Period I</th>
<th>Period II</th>
<th>Period III</th>
<th>MS narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occurrences per 1000 words</td>
<td>21,06</td>
<td>46,4</td>
<td>41,32</td>
<td>72</td>
</tr>
</tbody>
</table>

\(^{40}\) An important exception is Larm 1936, who studied the occurrences of -IN and den in the classical OS (= period I) and only briefly relates the discussions on the cliticization.
As can be seen, there is a significant rise in occurrences of -IN per 1000 words from period I to period II, with the values more than doubled. Surprisingly enough the values for period III are somewhat lower than for period II and neither reach the MS frequency.

There are considerable discrepancies between the texts in all periods. Table 10 shows the frequency of -IN per 1000 words for each of the studied texts separately, along with information on text lengths and the nominal number of -IN occurrences in each. As can be seen, only of the studied texts approach two MS values, namely the religious Linc NR in period III and SVM in period II. Otherwise, the three narratives in period III show very low numbers. We see similar discrepancies in period II between the lower values for narratives (comparable with the ones in period III) and religious texts.

<table>
<thead>
<tr>
<th>Period</th>
<th>Text</th>
<th>Length of the text (number of words)</th>
<th>Total number of occurrences</th>
<th>Per 1000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period I</td>
<td>AVL</td>
<td>15000</td>
<td>24</td>
<td>1,6</td>
</tr>
<tr>
<td></td>
<td>YVL</td>
<td>9200</td>
<td>111</td>
<td>12,06</td>
</tr>
<tr>
<td></td>
<td>OgL</td>
<td>10220</td>
<td>262</td>
<td>25,6</td>
</tr>
<tr>
<td></td>
<td>DL</td>
<td>10750</td>
<td>140</td>
<td>13,02</td>
</tr>
<tr>
<td></td>
<td>Bur</td>
<td>9770</td>
<td>280</td>
<td>28,65</td>
</tr>
<tr>
<td></td>
<td>Pent</td>
<td>10077</td>
<td>229</td>
<td>22,72</td>
</tr>
<tr>
<td>Period II</td>
<td>HML</td>
<td>10000</td>
<td>361</td>
<td>36,1</td>
</tr>
<tr>
<td></td>
<td>Jart</td>
<td>10623</td>
<td>671</td>
<td>63,2</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>16350</td>
<td>1112</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>KM</td>
<td>10940</td>
<td>186</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Bo</td>
<td>14300</td>
<td>668</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>JB</td>
<td>2400</td>
<td>113</td>
<td>47</td>
</tr>
<tr>
<td>Period III</td>
<td>PK</td>
<td>4728</td>
<td>103</td>
<td>21,78</td>
</tr>
<tr>
<td></td>
<td>Did</td>
<td>10945</td>
<td>357</td>
<td>32,62</td>
</tr>
<tr>
<td></td>
<td>Linc NR</td>
<td>9835</td>
<td>655</td>
<td>66,6</td>
</tr>
<tr>
<td></td>
<td>Linc JB</td>
<td>11088</td>
<td>541</td>
<td>48,8</td>
</tr>
<tr>
<td></td>
<td>Troj</td>
<td>9780</td>
<td>360</td>
<td>36,81</td>
</tr>
</tbody>
</table>

As can be seen, the oldest and linguistically most conservative AVL has, unsurprisingly, the lowest frequency of -IN. Of the three remaining legal texts, YVL
and DL show low frequencies as well, while OgL is close to the religious texts. This again is unsurprising since OgL is considered a legal text with relatively modern language. As we will see in 4.4 and 4.5, the lower frequencies coincide with lower level of grammaticalization of -IN in legal prose while the higher frequencies in religious prose coincide with a more advanced level of grammaticalization in period I.

In period II the text with significantly lower frequency of -IN is the narrative KM. KM is translated from French, not directly, but via an Old Norwegian translation of the text. Its conservative language (with respect to -IN) may at least partly depend, not only on the genre, but on the fact that it was translated from Norwegian, since another narrative from that period, SVM (based on a Latin original) shows the highest values for -IN among all texts in period II.

Finally, in period III, it is the narratives that show very low values in comparison with religious prose which reaches the MS frequencies of -IN.

From this overview we may conclude that the distribution of -IN is dependent on the genre and possibly on stylistic criteria. Religious prose shows high values of -IN in all periods. We will try to pinpoint the reasons for this regularity in 4.6.2, though we may state preliminarily that this is due in part to the inclusion of a large number of references to religiously ‘unique’ phenomena, such as ‘the truth’ and ‘the virtue’. On the other hand, one of the reasons for lower values of -IN in the narrative texts compared to the religious ones is a higher frequency of proper names in the former. Where PK names long lists of kings and princes (see example (1)), the religious texts usually operate with appellatives (2).

(1) Een hans son heet domalde (...) Domalde
one his son was-called Domalde Domalde

son heet domar (...) Domars son

son was-called Domar Domar-GEN son

heet attilla
was-called Atilla

(PK 228)

‘One of his sons was called Domalde (...). Domalde’s son was called Domar (...)
Domar’s son was called Atilla.’

(2) I them daghomen war en hölogher
in these days was a holy

biskopper i en stadh som heth
bishop in a town which was-called
Trajiktum (..) Oc nar biskoppen hört haffde
'In these days there was a holy bishop in a town called Trajiktum. (..) And when the bishop had heard..'.

In all periods there are large discrepancies between the texts, but nowhere more than in period I. Legal codices, which are the oldest texts available to us, show very low frequencies of -IN, whereas religious texts some 50 years younger show a surprisingly high -IN frequency and a variety of functions. One text in particular stands out, namely the oldest one, AVL. Since this text has only 24 instances of -IN and because of the importance of the text for our knowledge of OS, it will be discussed here in greater detail.

4.3. Äldre Västgötalagen (AVL)

AVL (1225) is the oldest Swedish manuscript written in the Latin alphabet and the proposed date for its formulation marks the beginning of the Old Swedish period. As it represents the oldest stage of language development that can be studied comprehensively, unlike the relatively short and formulaic runic inscriptions, it is one of the most extensively studied texts.

There are only a few examples of -IN in AVL (24 instances in a text 15000 words long). The absence of -IN in contexts that require it in Modern Swedish has been interpreted as either a consequence of the fact that the article was not yet formed by the time the law was composed (e.g. Delsing 2002), or that it was formed but belonged to the spoken rather than written register (Neckel 1924), or perhaps that it is typical of legal texts, even modern ones, to totally avoid articles and instead use bare nouns (Gunnarsson 1982). It has been noted that legal prose makes more use of bare nouns than other genres, e.g.

(3) Hund skall hållas kopplad på plats.
dog shall keep.PASS leash.PART on place

'A dog shall be kept on a leash in public places (lit: Dog shall be kept leashed in public place)'

(quoted after Dahl 2007:36)

Yet, even if a particular style is less prone to use articles, they must be present in their basic functions (like direct and indirect anaphora) if they are fully gram-
maticalized in a language. We may note that the uses in the example above are mainly non-specific, mass-like. With specific, countable referents, the articles are used even in the legal prose. The first article of the Swedish constitution (*regeringsformen*) reads:

(4)  Den offentliga makten utövas under lagarna.

‘The public power is exerted under the laws’.

(quoted after Dahl 2007:36)

Despite the low frequency, -IN is already found in AVL in a variety of functions, mainly textual.

### 4.3.1. Textual uses—direct anaphora in AVL

The anaphoric usage of -IN is obligatory in MS (unless a demonstrative is used). In AVL the anaphoric contexts are mostly unmarked, as in the following example.

(5)  Skil þem a bondæ oc presti kaller prester bondi
differ them on yeoman and priest says priest yeoman
tende oc bonde kuaþer ne vid byvþer firir sik
lit and yeoman says no by offers for self
tylder eþ at han tændi eigh þa er han
twelve oath that he lit not than if he
hauir svoret þa skal prester bötæ
has sworn than shall priest pay

‘If there is an argument between a priest and a yeoman (as to who lit the candles which started the fire in the church), and the yeoman denies having lit the candles and the priest says it was the yeoman who did it, the yeoman offers an oath of twelve that he did not light; if he has sworn then the priest shall pay the fine’.

When -IN is used, it seems to be used in order to actualize a referent mentioned some time ago and no longer easily accessible, as in (6). There are a number of syntagms between the first and the second mentions of the referent ‘man’. There is one use (underlined) of the pronoun *han* ‘he’ referring to ‘him’ in the middle of the text. There are however 3 other instances of *han*, which refer not to the man first
mentioned but to the one who injured him. Therefore perhaps the need to actualize the referent again by means of -IN.

(6) **Far** maþær sarr liggær i længi gangær i mælli
gets man wound lies in long goes in between

ok dör i þerri giuær arui sak þem ær
and dies in it gives heir charge he-DAT who

sargæþi ok kallær han banae vææ þa
wounded and calls he guilty be than

skal han væriæ sik maþ luctri hæræssnæmd æn
shall he defend himself with closed jury if

han dör innan nath ok iamlangæ gangær þær
he dies before night and year goes it

iuir þa bötæ ængtæ falls han at sak
longer then pays nothing falls he at charge

bötæ **manin** aptær
pays man-IN after (AVL BB:9)

‘If a man gets wounded and lies in bed long because of the wound but gets up and walks from time to time and finally dies, the heir presses charges against the one who hurt the man. The accused shall defend himself with a closed jury if the man died before a year and a night have passed. If a longer time has passed he pays nothing, if he is found guilty he should pay’.

It would be a simplification to say that all instances of anaphora marked by -IN are similar to (6). Example (7) shows that -IN could also be used in more straightforward cases. It must be noted though that this example comes from a later addition to the text.

(7) **Varþær** lekari sargafær þen sum maþ gihu gangar
be jester hurt he that with fiddle goes

allær maþ fíhlu far allær bambu þa skal
or with violin travels or drum than shall

kuighu taka otamæ ok flytæ up a bæsing
heifer take untame and lead up on hill

þa skal alt har af roppo rakæ ok sipæn
then shall all hair of tail shave and then
smyriæ þa skal hanum fa sko nysmurþæ
grease than shall him get shoes newly-greased

þa skal lekærin take quighuna vm
then shall jester-DEF take cow-DEF-ACC by

þæn goþa grip ok niutæ sum hundær græss
this good animal and enjoy as dog grass

(AVL LR)

‘If a jester gets hurt, the one that travels with fiddle, violin or drum, an untame cow should be taken up on a hill, the hair shaved off its tail and the tail greased and the jester be given a pair of newly greased shoes. Than shall the jester take the cow by the tail; if he can hold it, the good animal belongs to him and he should enjoy it as a dog enjoys grass’.

Note that the noun roppa ‘tail’ should be definite, in the first instance by reference to the cow, as a body part (see 4.4.3) and in the second at least as anaphora. It is not, however. On the other hand, the noun ‘cow’ is mentioned again with -IN and for the third time, by another term, where rather than -IN a preposed determiner and a weak form of the adjective is used. Apart from -IN, the preposed determiner den can also be used to mark anaphoric reference in AVL. These instances bring to mind the modern use of the pronoun denna/detta ‘this one’ and not of the determinative den/det. It is typical that the prior mention is easily accessible in the discourse, preferably occurring in the preceding syntagm.

(8) Læggær maþær manni fæ sit in til gætslu
puts man man cattle his in for safekeeping

þa ma þæt fæ eigh tapas af þem uiþ
than may this cattle not lost (be) of this by

takar
takes

(AVL RB:13)

‘If a man gives his cattle another man for safekeeping than the cattle may not be lost by the keeper’

Other similar examples include: kono, þa kono (AVL ÄB:8), hus, þem hvsum (AVL þB:5), bref, þem bref (AVL GB:8), fæ, þæt fæ (AVL FB:10), mark, þa mark (AVL aM:14), grip, þæn grip, þæn grip (AVL þB:13). These are different from those with -IN because the accessibility of the referent is higher. Though as we noted before, anaphoric reference in AVL is usually unmarked.
Indirect anaphora means that the referent has not been mentioned previously, but its familiarity is grounded in its connection with some other previously mentioned item or event. In example (9) the noun *sakinæ* ‘the charge’ refers to the charge of murder and its definiteness may be based on the use of the participle *dræpin* ‘killed’.

(9) Varþær  suþærman  dræpin  allær  ænskær  maþær  þa  skal
be  southerner  killed  or  Englishman  then  shall

böta  firi  marchum  fiurum  þem  sakinæ
pay  for  marks-DAT  four-DAT  the-DAT charge-DEF-ACC

sökir  ok  tvar  marchar  konongi
seeks  and  two marks  king-DAT

‘If a German is killed, or an Englishman, the crime is paid with four marks to the one pressing the charge and two to the king’.

An even more indirect link between the definite and the textual context is found in (10). The example presents the division of inheritance after the death of a man who has been married thrice with children from each marriage. If his wife is still alive, she may first take all that is left of her dowry. The remaining estate (*bono* ‘(of) the estate’) is to be divided into three parts, one per each group of heirs (children from the first, second and third marriage). Note that although the noun ‘estate’ is definite in connection with the yeoman who died (= his estate), it is not identical with it, but refers to ‘the estate after the dowry of the third wife has been removed from it’.

(10) Maþær  far  sær  aþalkono  gaþær  uiþ  barn  dör
man  gets  himself  wife  begets  by  child  dies

sv  fier  aþra  gætær  viþ  barn  far  hina  þriðiu
this  gets  another  begets  by  child  gets  that  third

þör  bonde  þa  en  kone  er  livændi  þa  skal
dies  peasant  than  if  woman  is  alive  than  shall

af  takæ  hemfylgh  sinæ  alt  þet  ær  vnöt
of  take  dowry  her  all  that  which  unused

ær  hun  ællær  hænær  börn  þa  skal  hin  ælsti
is  she  or  her  children  than  shall  that  oldest
‘If a man marries a woman and has a child with her, after her death marries again and fathers a child and marries for the third time and dies, leaving the woman, she or her children should retrieve her dowry—all of it that is unspoilt—then the children of the first marriage demand a part in the estate and should be awarded a third of it’.


Indirect anaphora is not marked obligatorily in AVL, as follows from example (11), where two identical contexts render (a) a structure lacking -IN, (b) a structure with -IN:

(11)  Viliæ þer bötaer take þa skal bötæ niu markær
want they fines take than shall pay nine marks

 arterböt oc tolf markær ættærbot Sæx mærkær
heir-fine and twelve marks family-fine six marks

skal (a) arui bøta sex mærkær skal (b) ættin bøta.
shall heir pay six marks shall family pay

(1V aM:1)

‘If they accept fines (as a reparation after a relative was killed) than the heir shall pay nine marks and the family six’.

This variation is found when one referent is a countable noun in the singular, whereas the other either a countable noun in the plural or a mass noun (*arvi* ‘heir. pl’ and *ätten* ‘family-sg.def’).

Although it is claimed in literature that indirect anaphora is a context reserved for definite articles and possessives in which demonstratives cannot appear (see Chapter 2.5.1), there are occasional examples which include a demonstrative and whose most probable interpretation is indirect anaphora.

(12)  Uærþær maþer dræpin ok af daghum takin þa skal
be man killed and of day taken then shall

 uighi a þingi lyse oc frafal aruingiae
manslaughter on ting announce and loss heir

41 The provisions are made here for the situation when both the victim and the killer are dead and the matter of fines for the manslaughter is to be resolved between their respective families.
‘If a man is killed than (the killer) shall make the killing known on the ting and inform
the heirs, so on the second ting. But on the third (the heir) must bring charges or he
forfeits his case. Then shall the killer come to ting and stand outside and send men to
the ting to ask for peace. Tingsmen must promise him to come to the ting. He shall
admit this accusation (of killing).’

A special case of indirect anaphora is inalienable possession. It is usually
marked with the definite article in MS and occasionally by a possessive (see
Chapter 1.5). In AVL, however, it is regularly unmarked.

(13) Værdær maþær huggin i houod ma seiæ hinnu
is man hit in head may see membrane

allæ båþi uþa bötæ firi markum III
call both accident pay for marks three (AVL VS:1)

‘If a man is hit on the head so that it splits (so that membrane may be seen) and both
call it accidental, (he who hit) shall pay 3 marks in fines’.

As well as direct anaphora, indirect anaphora is also typically unmarked in
AVL. Example (14) contains what is definitely a link between the first and the
second referent, though not inalienable possession—yet bonda ‘yeoman’ is un-
marked.

---

42 In the case of manslaughter, the killer could avoid a feud by confessing the crime on ting and
asking the victim’s family for peace.
(14)  thræl  far  væþæ  sar  ma  eighi  firi  bonda
     thrall    gets    accidental    injury    may    not    for    yeoman
     byaergas  liggær  attæ  daghæ  pa  skal  böte  örtogh
     bring-IN    lies    eight    days    then    shall    pay    örtogh

     ‘If a thrall is injured accidentally, may not bring in (hay) for the yeoman (= his owner)
     and lies down for eight days, than shall (he who injured) pay an örtogh fine’.

(15a)  A  dözsdæghi  ma  ikki  fra  aruæ  giuæ  at
     on    deathday    may    not    from    heir    give    by
     laghmæli  num  arvi  quædær  sialvær  ia  viþr
     law    unless    heir    says    himself    yes    to

     ‘On the day of death one may not lawfully give away (property) from the heir unless
     the heir accepts it’.

(15b)  MS translation:
     På  dödssdagen  må  man  icke  enligt  lag  giva  bort  från  arvingen,  om  ej  arvingen  själv
     säger  ja  därtill.

           (Holmbäck  &  Wessén  1979:77)

4.3.3. Non-textual functions in AVL

Even though the textual uses of -IN are only sporadic in AVL, the form does
occur in non-textual contexts. Although the names of legal powers are always bare
nouns (kunungr), there is one, used twice, that in both instances is definite, namely
pave ‘pope’.

(16)  allum  slikum  malum  skal  af  landi  skiptæ  mæþ  brevi
     all  such  case  shall  of  country  part  with  letter
     til  pavanum  bref  takeæ
     to  pope-DEF-DAT  letter  take    (AVL  GB:8)

     ‘All such cases are beyond the local jurisdiction and (the perpetrator) shall be sent
     away from the country with a letter to the Pope in Rome and a letter from the Pope
     (containing absolution) be taken back’.
Apart from these two instances, no examples of -IN in unique or generic function can be found in AVL. Uniques are unmarked, as in (17).

(17) Tialdru stenæ skal tua i iorþ grawæ þriþiæ boundary stones shall two in earth dug third
    a læggjæ þer skulu vitni bæræ ar on lie they shall testimony bear which
    i iorþ liggiæ in earth lie

‘Boundary stones shall be dug two in the earth and third placed on top. These bear witness (to where the boundary is located) which lie in the earth’.

It is not always straightforward which uses of NPs are generic—since this is a legal text it is reasonable to state that virtually all are. Therefore all uses of man and bonde could be translated as en man, en bonde meaning any man, any yeoman. One similar instance (18) is given below.

(18) Konongær vil sær kono biptia Ær þæt vtæn king wants REFL wife bid is this without
    konongrikiz þa skal mæn sinæ latæ faræ ok kingdom then shall men his let travel and
    ærændi sit vrakæ ok fæst takæ þa skal errand his execute and engagement take then skal
    konongær brudfærþ gen gæræ þa skal king bride-journey toward make then skal
    konongær gen fara king towards travel

‘If the king wishes to marry and the prospective bride comes from abroad, he shall send his men to present his case and secure the engagement. Then the king shall make a journey to meet his bride’.

While it would be natural in MS to use the indefinite article here, it is not yet developed in OS and a bare noun is used instead. There are, however, three instances which are difficult to classify, since these too would receive indefinite marking in MS, but in AVL appear with -IN. One is given in (19).

(19) Havir konæ hor giort ok givær bondæ kono sinni has woman adultery done and gives husband wife his
sak þa væri sik mæþ siv mannum af charge than defends herself with seven men of

næmdinni.
jury-DEF

‘If a woman has committed adultery and the husband charges her with it, she should defend her case with the testimony of seven men of the jury’

We shall return to this surprising use of -IN in 4.6.1.

Even with very few instances of -IN we may still describe its use in the oldest Swedish text as mainly textual though largely absent in short-distance direct anaphora, instead actualizing a referent mentioned some time ago. In none of the functions is the use of -IN obligatory, yet it is found with some unique referents.

4.3.4. Distribution of -IN forms in AVL

The distribution of -IN forms in AVL across genders, cases and numbers is given in Table 11. As can be seen, -IN in AVL is well represented in the singular for all genders and cases, with a slight preference for masculine nouns in the nominative and accusative.

Table 11. The distribution of -IN in AVL by number, gender and case

<table>
<thead>
<tr>
<th>Number</th>
<th>Gender</th>
<th>nom</th>
<th>gen</th>
<th>dat</th>
<th>acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg</td>
<td>M</td>
<td>saksökæn kirkuldrotten lekarin saksökæn bondin ættin ættin</td>
<td>pavaens præstins</td>
<td>pavanum arwumn* saknæ næmdinni</td>
<td>banæn saksöchenden manin skañæ svnudaghin quighuna skyldpinæ</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the plural only one certain instance was found, of the neutral noun rike ‘kingdom’, given in (20).

(20) þer tolf sættu sæx stenæ mællir rikiænnæ
they twelve sat six stones between kingdom-PL.DEF.GEN

‘These twelve sat six stones between the kingdoms’.

*see Excursus below for a discussion of this particular form
4.3.5. Excursus: on the form ‘arwumn’ in AVL and YVL

In Table 11 we have classified the form *arwumn*, found in AVL ÄB, as singular masculine, dative case. The context in which it appears is given in (21).

(21) Dör landboe firi fardaghæ þa æru tver løter
    dies tenant before removal-day3 then are two parts
loll af bolæ hans. Dör kona hans þa ær
free of estate his dies wife his then is
þriþiunger lös af bolæ hans. Er rughí sait
third free of estate his Is rye sown
a bole þa skal þen ær bol a gialda
on estate then shall this which estate owns pay
arwumn iammarghar skæppur sum a ær sait.
heir even measures which has been sown

(AVL ÄB:24)

‘If a tenant dies before his contract is concluded then two parts (thirds) of the estate are free. If the tenant’s wife dies, one third of the estate is free. If rye is sown then the owner of the estate shall give the heir the amount that was sown.’

This form has been the subject of many speculations and is probably one of the most frequently discussed in literature. It is the dative form, going back to:

(22) arv(i)-om-IN-om
    heir-DAT.PL-DEF-DAT
    ‘(with) the heirs’

Such lengthy forms are rarely found, instead they are simplified, the simplification different in the West Nordic (Icelandic, Norwegian) and the East Nordic (Swedish, Danish).

(23) WN arw-om-IN-om → arwunum
    EN arw-om-IN-om → arwomen

Noreen, Pipping and Sjöros complete the form to *arvunum*, which is dative plural (see Larm 1936:24). However, this is the form as found in West Nordic (fornvästnordisk), otherwise not accounted for in the Swedish literature, so this interpretation is quite unlikely, as observed by Larm (1936:24, footnote). It is more plausible that the form should be completed to *arvumen*, masculine dative in the

---

33 OS *fardagher* was the day when workers were released from their employment and could enter a new one; the meaning intended here is close to ‘within the valid time of the contract’.
singular, a more likely one in the context. This is proposed in Cederschiöld 1898, a proposition also motivated by a parallel in YVL.

(24) Dör lanbec firi fardaghæ þa æru tuer löter
  dies tenant for removal-day then are two parts
  løser af bole hans ær rughi sat a bole
  free of estate his if rye sowed on estate
  þa skal þen ær bol a giælpæ aruænum
  then shall DET which estate has repay heir-DAT-DEF
  æmmanger skjæpur sum ær a sat
  even measure which been has sowed (YVL ÆB:33)

'If a tenant dies before the contract is concluded then two parts are free of the estate. If the rye is sown than the owner of the dwelling shall repay the heir the same amount that had been sown'.

Because of this parallel and the unlikely reconstruction of the form in its West Nordic variety, it is classified here as masculine dative singular. The reconstruction is further strengthened by the fact that the definite dative plural is a relatively late addition in Swedish (Ståhle 1980, Skrzypek 2005).

4.4. The development of textual functions

4.4.1. Direct anaphora

Direct anaphora is considered the original context of demonstrative use that ultimately results in the formation of a definite article (see Chapter 2.5.1 for discussion). One would therefore expect that -IN will be well represented here. As we have seen in 4.3.1, there are only few instances of anaphoric reference with -IN in AVL. A similar situation is found in its continuant, YVL. This partly depends on the fact that large excerpts of YVL are identical with AVL (Holmbäck and Wessén 1979:xliii, Skrzypek 2009). But other legal texts from period I: OgL and DL, are also economical with -IN in direct anaphoric functions. The situation changes first with religious prose in period I, Bur and Pent.

Generally in period I there is a great amount of variation between -IN and demonstratives in indirect anaphoric contexts. The general rule seems to be that the more easily accessible the referent is, the less likely it is to be marked with -IN
(see examples in 4.3.1). Accessibility is understood here as referential distance (Givón 1983:13f.)—the gap between the previous occurrence in the discourse and its current occurrence, measured by the number of clauses to the left of the clause in question. The minimal distance and maximum accessibility are thus equal to only 1 clause—when first and second mention appear in neighbouring clauses. What Givón terms ‘presence in the register’ (Givón 1983:14) may be expressed by anaphoric pronouns, definite descriptions, etc., but also by zero anaphora.

A system similar to MS is found in the OS religious prose. Consider the following example from Bur:

(25) En vælburin ungar suen forlæt værudena ok
one well-born young boy left world-DEF and

folgþe andream: hans frændar vildo þom baþa
followed Andreas his relatives wanted them both

inne brænna ok tændo brand iui þera herberge
in burn and started fire over their shelter

Smasuennen slaekte balet mz litlo vatne þa
young man-DEF quenched fire with little water then

vildo þe kliva i huset. ok wrþo alle
wanted they step in house-DEF and were all

iæm skyt blinde
suddenly blind

‘A young well-born boy left the world and followed Andreas. His relatives wanted to burn them both and started a fire over their shelter. The young boy put out the flame with some water. Then they wanted to step into the house and (as they did so) they all turned suddenly blind’.

In this example, there are three instances of textual definiteness: smasuennen ‘the young man’, balet ‘the fire’ and huset ‘the house’. Apart from the first, which is a repetition of a similar term used to present the new referent, the others are anaphoric but do not make use of the same term (see the underlined brand and herberge).

However, not even in Bur do we find the marking consistent or obligatory. Instead, there is variation between -IN, bare nouns and demonstratives, even within one paragraph.

(26) En tima þær iohannes prædicæpe fik han
one time there Johannes preached got he
se en ungan suen vaenan ok þo vildan
see a young boy beautiful and though wild

johannes fik biskope suenen ii hand (..) Johannes
Johannes got bishop-DAT boy-DEF in hand Johannes

kom atar tel staþen ok sporþe at
came back to town-DEF and asked about

suenenom biseopen sagþe hanom vm suenen
boy-DEF-DAT bishop-DEF said him about boy-DEF

som sant var Johannes gaf biskope skuld
which true was Johannes gave bishop-DAT guilt

for vangømo
for negligence (Bur 156)

One time as Johannes was preaching he saw a young boy, beautiful but wild. Johannes gave a/the bishop charge of the boy. (..) Johannes returned to the town and asked about the boy. The bishop told him as true was. Johannes blamed the bishop for negligence’.

MS would have the second and third instance of biskop ‘bishop’ definite, by anaphoric reference. The first might be definite, meaning ‘the bishop under whose jurisdiction the story takes place’ or indefinite ‘a certain bishop’. In OS, what we see can at best be described as inconsistency in the definite marking. It is only natural that in a transit phase, before the obligatorification of the definite article in anaphoric contexts, we would expect some instances to be definite and some not. However, the question is whether the variation we see in Bur is free or if there are some restraints on the use of the definite.

It is the third instance in example (25), the bare noun, that is difficult to account for. A possible interpretation takes into account the distance between the first and the second mention (7 syntagms) and between the second and third (1, as they occur in neighbouring syntagms). This example further strengthens the point made earlier that although anaphora may be considered the original context where the grammaticalization of the definite article is initiated, it is by no means as simple as obligatorification of -IN in each anaphoric context. Rather, these are divided into less and more accessible anaphora and, to begin with, -IN is only found in the former, while in the latter either a bare noun or a demonstrative is used. A similar example is given in (26):

(26) guz ængel teþes andree ok baþ han fara tel
god’s angel shown Andreas and bade him travel to
The development of textual functions

burgundiam  ok  hialpa  maþeo  andreas  sagþe  sik  eigh
Burgundy  and  help  Matthew  Andreas  said  self  not

vita  væghen  en  ængel  baþ  han  ganga  tel
know  way-DEF  but  angel  bade  him  go  to

strand  ok  fara  i  førsta  skip  han  funne
beach  and  travel  in  first  ship  he  find

Engelen  var  hans  leþsaghare
Angel-DEF  was  his  guide

‘God’s angel has appeared to Andreas and asked him to travel to Burgundy and help Matthew. Andreas said he did not know the way but the angel asked him to go to the beach and take the first ship he would find there. The angel was his guide’.

The variation between -IN and bare nouns in anaphoric contexts has been studied in Leiss 2000. Her study concerns Old Icelandic (and Old High German and Gothic), but the results and hypotheses should be valid for its close relation, Old Swedish. Leiss quotes an example much discussed in literature and used by several authors as an illustration of the instability of the use of -IN. It is similar to the Bur example quoted previously in its use of bare noun in direct anaphora. The example is given in (27).

(27) þá  verþr  hann  varr  viþ  gripungenn  ok  reið
then  became  he  aware  of  ox-DEF  and  rode

til  ok  vill  drepa  hann  gripungr  snýr
to  and  wants  kill  him  ox  turns

í móti  ok  kom  konungr  lagi  á  hann  ok
towards  and  came  king  stab  on  him  and

skar  ór  spjótit  gripungr  stakk
cut  out  lance-DEF  ox  stuck

hornonom  í  síþo  hestinum  svá  at
horn-DEF.PL.DAT  in  side  horse-DEF.DAT  so  that

hann  fell  þegar  flatr  ok  svá  konungr  þá
he  fell  there  flat  and  so  king  then

hljóp  konungr  á  føtr  ok  vill  bregða  sverði
jumped  king  on  feet  and  wants  draw  sword
(King Egil was a great hunter, and often rode into the forest to chase wild animals. Once he rode out with his men to hunt in the forest. The king had traced an animal a long while, and followed it in the forest, separated from all his men.) He observed at last that it was the bull, and rode up to it to kill it. The bull turned round suddenly, and the king struck him with his spear; but it tore itself out of the wound. The bull now struck his horn in the side of the horse, so that he instantly fell flat on the earth with the king. The king sprang up, and was drawing his sword, when the bull struck his horns right into the king’s breast. The king’s men then came up and killed the bull.


This example from Heimskringla (ca 1230) was to the author’s knowledge first cited by Nygaard (1905) in his discussion of the irregularity of the definite marking in Old Norse.

Leiss proposes to take the position of the NP in the information structure of the sentence into account. Thus griðungr ‘ox’ appears as a bare noun in the initial subject position (theme) but in the definite form in the object position (rheme). She claims that the clause-initial position, the ‘thematic position’ in her terminology, is by default definite (so-called definiteness effect) and need not be marked with an article. The non-initial, rhematic position in which we find objects is, by default, indefinite (contains new information) and if the object is definite, it needs to be marked as such. Thus the hypothesis is that Old Norse was a hypo-determinative language, which marked definites only in the rhematic position.

Though this explanation is very appealing, we see immediately that it does not fit in with the OS data. In the example from Bur (example (25)) we see the reverse: definite marking in the thematic position and none in the rhematic.

---

\[\text{See Chapter 5, on the role of EN as a presentative marker, for a discussion of the clause-initial position as introduction of new text-topics. Contrary to the views presented in Leiss 2000, the initial position seems to have been used to introduce new, topical information, at least in periods I and II of OS.}\]
In Møller’s interpretation of the same example the unmarked form is used when the animal is the agent (et handlende Væsen, Møller 1945:36) like the unmarked konungr ‘king’. He quotes another example though, where the ox is clearly agentive as well though it appears as the subject in a subordinate clause and is marked. In its non-agentive role however it is unmarked and referred to by a different term (Møller 1945:36).

Example (25) from Bur 156 with which we started our discussion introduced a new referent biskop ‘bishop’ in bare noun form, its second mention was marked with -IN and the third and last mention was a bare noun again. This variation, or inconsistency from the MS point of view, is most probably an indication of the language being in a transit phase—from one where anaphora is unmarked to one where anaphora is obligatorily marked. It is to be noted that a similar variation is found elsewhere in Bur, and usually takes the same form: introduction by a bare noun, marked forms and finally unmarked again. This is illustrated in examples (28) and (29). In (28) the use of bare nouns may depend on a different description of the referent, in (29) however the first and last descriptions are identical.

(28) En syndogh konan var mz ens mandrapara
one sinful woman was with one-GEN killer-GEN
barne ok qualdes ii barnsiæng (..). ok løste
child and suffered in childbirth and liberated
salogha kono af qualum.
blessed woman of suffering

‘A sinful woman was pregnant by a killer and suffered in childbirth (..) and freed blessed woman from pain’.

(29) Þa kom goþar gæstar andreas. (..) ok bæd in
then came good guest Andreas (..) and bade in
lata goþan gæst
let good guest

‘Then came Andreas, a good guest (..) and bade welcome the good guest’

Though OS material does not support Leiss’ hypothesis, we nevertheless see that textual organization of given vs. new information may be an important factor in the rise of the articles. This will be seen more clearly in the next chapter discussing the rise of the indefinite article.

45 Several mentions as gestren ‘the guest’ in between.
The textual evidence is otherwise of a system only being formed, as example (30) shows, where it is only the fourth mention of the same referent *kirkia* ‘church’ that is marked with -IN, the others appearing as bare nouns.

(30) Uærther *kirkia* brutin oc *maessufat* wt stolen that
is church broken and vestment out stolen that
ær nithinxuærk oc niu marka sak *kirkyu* siælfue
is outrageous-deed and nine marks charge church self
oc suæ hæraþe oc suæ kononge. Uarther han
and so county and so king is he
fangen i *kirkyu* ælla annarstath maeth thy
cought in church or elsewhere with this
the stalo aff *kirkyunne* tha late lijf fore
they stole of church then let life for
(YVL KB:11)

‘If a church is broken into and vestments (liturgical garments) are stolen, it is an outrageous deed and the charge of nine marks payable to the church itself, the county and the king. If he (=the thief) is caught in the church or elsewhere with what they stole from the church, he must give his life’.

The laws clearly represent the oldest stage of development, with direct anaphoric marking very irregular and sporadic. Bur is more advanced with most instances of direct anaphora marked and the unmarked instances occurring in curiously uniform patterns with referents introduced by means of bare nouns, referred to by definite nouns and finally as bare nouns again. The youngest of period I texts, Pent, only rarely leaves direct anaphora unmarked. We find though a story of Abraham sending an emissary to find a wife for his son Isaak. The emissary is described several times as *svenen* ‘the boy’, *mennen* ‘the man’ and as *sændebudin* ‘the emissary’. In one instance the term *sændeman* ‘envoy’ is used, in the middle of the story, and this time as a bare noun (Pent 158). Whether the form depends on the use of a slightly different term or not, the anaphoric chain is broken here as in other cases we saw previously.

On the other hand, we find in Pent an interesting example of how -IN is used in anaphoric contexts. In the following passage jealous brothers of Joseph are plotting how to get rid of him and win their father’s favour in his stead. They propose to kill him and hide the body by throwing it into some old well. The proposal is opposed by Ruben, who, not wishing Joseph to die, suggests that he be thrown in the well alive and left to die, while secretly planning to rescue him. Consider how the passage is formed:
We all have one will so we shall kill him and throw him, dead, down in some well (...). Ruben, his oldest brother, overheard these plans and put all his mind to saving the boy and spoke to them. Do not destroy him and spill his blood with a sword but throw him alive in the well and let him die there himself.

The second mention of the noun brunn ‘well’ is marked, though there is probably no specific well that is meant.

In period II, marking of the anaphoric reference seems to be nigh-on obligatory. There are still occasional exceptions, though. A striking example is given in (32). It shows a fragment from the beginning of the tale, which describes how king Charles became Charlemagne—when he took the seat reserved for God in the church of Our Father (Pater Noster), with his twelve companions filling the apostles’ places. The king is introduced in the definite form, and referred to by pronoun han ‘he’ as well as his name and title Karl konung46 ‘Charles king’, a structure where the noun normally lacks an article. Two other references, however, are made by an article-less form; again the story is finished with a definite.

46 Note that the nominative -r has disappeared by the end of period I.
It is difficult to interpret this variation in form: in all instances konung ‘king’ or konungin ‘the king’ is the nominative form, fulfilling the role of a subject (once in a subordinate clause). The only regular variation may be noted in regard to the position of the noun before the main verb of the main clause (definite) or after the verb (no article). The variation is similar to one found in Bur (see example (25)), which is older than KM by some 100 years. This is the only fragment in period II where this variation is found.

Anaphoric reference may be made in OS by either the suffix -IN or the preposed demonstratives den or dessa. In KM it is mostly the suffix that is employed for this purpose, sometimes reinforced by a demonstrative (see examples 33 and 34 and Table 12). Though the numbers are high for definites and low for demonstratives, the variation in their use is still similar to that seen in Äldre Västgötalagen (see 4.3.1)—when the accessibility is low (measured in a number of intervening syn-
tagms), -IN is more likely, as with the noun *kirkiona* ‘church’ in (32) above which takes up a referent mentioned much earlier in the text (9 syntagms, 136 words); when the referent is easily accessible, the demonstrative may be used (33).

A comment is due on the order of the elements in the phrases in bold in (33). It suggests that there may be a difference between postposition of the demonstrative (*kirkio the som* ‘church this which’)—introducing a subordinate clause and a preposition (*the kirkio* ‘this church’) where a demonstrative function is present. The short-distance anaphora is marked by the demonstrative *den*; the lack of stylistic variation—both sentences 2 and 3 start in exactly the same way, with ‘in this church’—may seem disappointing in a literary text. We may note, however, that sentences 2 and 3 are not connected so that they form one paragraph, rather, sentence 2 belongs together with 1, whereas 3 opens a new paragraph. This may be reflected in the use of a demonstrative rather than a definite article, a tendency well-documented in narrative texts elsewhere.

Furthermore, it may occasionally happen in the course of the text that a few mentions of a definite referent occur here and there with a demonstrative, especially at the beginning of paragraphs or on the very last mention of the same referent, especially at paragraph boundaries. One illustration of this type of sequence is found in a short story called *Golden Ox Tale* (Walls and Walls, 1983), in which we find 26 occurrences of the word *lake*.

The sequence of definite articles vs demonstratives is entirely typical of narrative texts:

A long time ago, West Lake was called Golden Ox Lake.
At that time, this lake was just one vast expanse of plain water. When the farming was done, people would go to the lake.

At the bottom of this lake there lived a golden ox.

The 19 subsequent mentions of the lake all occur with a definite article. (Cyr 1993: 206-7)

It should be noted that a reference is made to the church once more in the text, by means of -IN, after 9 syntagms, which validates the claim that a demonstrative is preferred in cases of short-distance anaphora while -IN dominates in long-distance anaphora.

(34) oc gik til kirkiona oc konungin stodh
and went to church-DEF and king-DEF stood
op oc negh honum
up and bowed him

'And he went to the church and the king stood up and bowed to him'.

This pattern of usage: initial mention of an entity, a second reference to it by means of den or dessi, followed by a further reference to the entity by means of -IN, is known from the West Nordic literature and has been discussed by, for e.g., Hodler (1954:102) and Johnsen (1976:80), who identify the use of the demonstrative with increased emphasis. Without necessarily agreeing with this interpretation we may bear in mind that KM was in fact translated from Old Norwegian and most likely retains some of its stylistic traits.

The demonstrative is employed whenever the distance between the first and second mention of the referent is short, even if the noun is definite in the second mention, see (35) and (36) below.

(35) Nw kom k. m. k. til mykla gardh oc
now came Charlemagne to Miklagaard and
hitte konungin sitia i sinom vrtagardhe Gardhen
found king-DEF sit in his herbgarden garden-DEF
then war alder aff gulle
DEM was all of gold

'Now Charlemagne came to Istanbul (Miklagård) where he found the king sitting in the herbal garden. This garden was wholly made of gold'.

(36) oc en stolpe stodh vnder hallene ther stodh
and one pillar stood under hall-DEF there stood
And there was one pillar under the hall, on which the whole hall rested. Round the pillar there were a hundred small pillars.

The function of -IN to actualize a referent mentioned a long time ago is confirmed in SVM as well. Consider the following example, where mörkastoffwone ‘the dark room (= prison cell)’ picks up a referent introduced 30 (!) syntagms earlier, with a whole different story intervening between the first and the second mention\textsuperscript{47}.

\begin{verbatim}(37) oc lät han swa läggia i mörkastoffwo and let him so lie in darkroom
(38) Mit i lunden war et stort trä mz middle in copse-DEF was a large tree with
myken frwkt (...) En fähirdhe kom thit mz fä oc
many fruit a shepherd came there with cattle and
saw that DEM tree was full with ripe fruit
oc foör op i träädh (...) and went up in tree
Walbassen (...) hogh trädh mz tannomen (...) Tha wardh
boar-DEF struck tree with teeth-DAT-DEF then became
swinith spakare oc aat fast Tha thz war mät.
pig-DEF quieter and ate quick when it was full
tha lagdhis thz nidher widh träädh then lay it down by tree
\end{verbatim}

\textsuperscript{47} It is possible that the superior story was written first and the additional small stories added later, thus the distance between the first and second mention would in fact be much shorter.
‘In the middle of the copse there was a large tree with much fruit. A shepherd came by with his flock and saw that this tree was full of the ripe fruit and climbed up on it. The boar (living in the copse) struck at the tree with its teeth. Then the beast quietened down and ate quickly. When it was full it lay down by the tree’.

Table 12. Direct anaphora in KM* (p. 250-258)

<table>
<thead>
<tr>
<th>Formal marking of anaphora</th>
<th>DEM den</th>
<th>DEF -en</th>
<th>DEM den + DEF -en</th>
<th>DEM dessa</th>
<th>Zero</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of instances</td>
<td>3</td>
<td>18</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

*Not considered: second and following mentions of a referent unique or any that was introduced by the definite form, such as patriarchen ‘patriarch-DEF’. Since it is introduced by the definite, the anaphoric usage plays no part in allotting the definite form.

In Table 12, all possible markings of anaphora are listed, together with their frequencies. As can be seen, example (32) discussed previously is the only one in KM with bare noun in anaphoric context. The dominating type of marking is by the definite article, occasionally alongside the demonstrative and the usage greatly resembles MS. Generally, it can be noted that in period II it has become obligatory to mark anaphoric reference, though -IN is not the sole exponent of that marking.

In period III it is obligatory to mark both direct and indirect anaphora and bare nouns are not found in this context. However, just as in period II, the marking need not be or include -IN. In fact, in direct anaphoric contexts -IN is often reinforced with a demonstrative.

(39) oc giordhe ther til som eth litet altara gudhi and made there to like a little altar god
til hedhers, oc kalladhe thz wars herra kirkio,
to glory and called it our lord-GEN church
offrandhe oppa thz altaret sacrificing upon this alter-DEF (Linc NR 279)

‘And (she) made a form of a little alter for the glory of God and called it Our Lord’s Church and made her offers on this altar’.

In MS, such use would put an emphasis on the second mention—whether that was the case in OS we cannot know, but the high figures challenge such hypothesis—the point of an emphasis is to single out a referent, but DET N-IN is used so frequently in period III, it does not appear to serve this purpose.
Table 13 presents types of anaphoric marking (by direct anaphora only) in Linc NR. The types given in bold are the same as MS types (den N-IN requires stress on den). As can be seen, in Linc NR, denna is always followed by -IN which is still possible in Swedish dialects today.

The anaphoric context no longer allows bare nouns and even if the noun is modified by a determiner like den or denna the definite suffix is used. One prolific type in Linc NR that is not used in MS for anaphoric functions is ADJ-a N-en. It is not used in anaphoric contexts in MS, but is limited to phrases denoting proper names or concepts closely resembling proper names, see examples below:

(40) Vita huset
white-DEF house-DEF
‘The White House (in Washington)’

(41) det vita huset
DET white-DEF house-DEF
‘the white house (you know which)’

In Linc NR the majority of ADJ-a N-IN instances are references made to the heroine of the story, hálga iomfrun ‘holy virgin-DEF’, sometimes followed by her name, Amalbergha. They may be seen as repetitive descriptions of her, equal to using her name alone, and thus nearing the MS usage of this type of defNP.

Low values of -IN in Troj are surprising. It is the youngest text, therefore we would expect anaphoric marking to be well-established. It is also a historical narrative, where we would expect higher values of anaphora than seen in the statistical results. It is reasonable to assume therefore that anaphora in Troj may be marked by means other than -IN. Similar to other historical narratives such as PK in period III, these partly include the frequent use of proper names rather than def-NPs such as kungen ‘the king’ etc. There is also a competing marking by a relative clause introduced by hwilken ‘which’ instead of a second mention with -IN.

---

48 The whole legend, ca 9835 words.
49 See also Chapter 1.5.
50 The modifying adjective here, hælagher ‘holy’, tends to be followed by unmarked noun, as in the two lexicalized phrases, den helige ande ‘the holy Spirit’ (in NT 1981 changed to den heliga anden) and den heliga skrift ‘the holy writing = the Bible’. For more details concerning it, see Hirvonen 1997.
When the dragon saw him coming, he made a terrible sound with much hiss and squeak, which sound reverberated in the air.

Such use is quite foreign to MS, although it is perfectly acceptable if the noun were omitted:

When the dragon saw him come, it let hear a terrible sound, which reverberated in the air.

In Table 14 follows an overview of the types of anaphoric marking. Relative clauses introduced by *hwilken* ‘which’ + noun are listed in the last column.

<table>
<thead>
<tr>
<th>Type of marking</th>
<th>-IN</th>
<th>den N-IN</th>
<th>den ADJ-a N-IN</th>
<th>denna N-IN</th>
<th>denna ADJ-a N-IN</th>
<th>den N</th>
<th>denna N</th>
<th>ADJ-a N-IN</th>
<th>hwilken N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of instances</td>
<td>124</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>-</td>
<td>8</td>
<td>14</td>
</tr>
</tbody>
</table>

Anaphora in Troj is marked overwhelmingly by -IN, which occurs in most NPs, with no instances of *den N* or *denna N*. Parallel to Linc NR, a combination of weak adjective and noun modified with -IN is used in Troj in this context, exclusively marking the phrase *gyllene ullen* ‘the golden hide’, which is the object of Jason’s quest in the first part of the story. It may also be referred to as:

---

51 Pages 1-39, 9870 words.
(44) then gyllene wllen (Troj 2) the golden hide-DEF
    then gyllene wll (Troj 13) the golden hide
    wllen (Troj 17) hide-DEF
    than gul wllen oc dyra (Troj 3) the gold hide and dear/expensive

Even in this variation, the phrase gyllene ullen is by far the most frequent, just as heliga jungfrun was in Linc NR. Since these are the main themes of the stories, the phrases may be considered similar to their role in MS, proper name-like.

4.4.2. Indirect anaphora

As discussed in Chapter 2.5.1, indirect anaphora is the usage that links the anaphoric, stemming from the original demonstrative deixis, and the ‘larger situational uses’, including unique referents. It is therefore a complex context, with both straightforward and difficult examples. In this context we find variation between the definite and the possessive rather than between the definite and the demonstrative as found in the direct anaphoric contexts. -IN is used in contexts where the possessive would be equally natural, as in the following example:

(45) Ƿa sanctus iohanes disputære þesse skæl þa
    when saint John disputed these reasons then
    forþes en doþar tel graf þrætighi daghum
    lead.PASS EN dead to grave thirty days
    før giptar moþeren ok ænkian vini
    before married mother-DEF and widow-DEF friends
    ok frændar fiollo for kne iohannis ok bedes þøm
    and relatives fell for knee John and bade DET
    dœþa lifes naþer
    dead life-GEN grace
    (Bur 154)

‘When John was discussing these matters, a dead (man) was being buried, one that was married thirty days earlier. The mother and the widow, friends and relatives, fell on their knees before John and bade him for the grace of life for the dead’

-IN is found in indirect anaphoric contexts as early as in AVL (see 4.3.2) and in other legal texts, though its use is far from regular. Note also that in the Old Norse example given in (27), though -IN was not regularly employed in
the direct anaphoric contexts, it was used to marked indirect anaphora such as spjótit ‘the lance’ and hestinum ‘the horse’, definite because they belonged to one of the main characters of the story, the king. It shows that obligatorification in a certain context is not necessary for the item to grammaticalize further, but the grammaticalization can proceed even though the item is not yet obligatory in other contexts.

(46) Wærther mather  i  kirkyugarþe  dræpin  ælla  sarthather
de be man in churchyard killed or injured

göre  sokn  natta  gengærþ  oc  þæn  gæringena
make parish night provision and this deed-ACC-DEF

gyorrhøe  gæelde  ater  soknamannum  fult  fore  kosten
did pay back parishioners fully for cost-ACC-DEF

(47) Nv  kan  præster  dō  a  þæt  ar  biskuper
now can priest-NOM die on this year bishop

skal  ey  gengærþ  hafua  tha  taki  alt
shall not provision have then take all

‘If a man is killed or injured in a churchyard, than the parish ought to give provision and he who is guilty of the deed shall pay back the parishioners’.

The use of -IN on gæringen =the deed’ = the killing is induced here by the VP varder drapin and kosten ‘the cost’ = the cost of providing it by earlier gengerd ‘provision offered to king, bishop or king’s guard on their travels’. Both marked nouns have another nominal element as their trigger (anchor, Fraurud 1990:423). However, the trigger for the defNP may also be a VP denoting some action, for which the following defNP is an agent, a patient, an instrument, a result etc. Such a relation is found in the following example, where the VP kan præster dō ‘can priest die’ triggers the marked forms alt gozet ‘all the goods (he left)’ and arfuin ‘the heir’.

(47) Nv  kan  præster  dō  a  þæt  ar  biskuper
now can priest-NOM die on this year bishop

skal  ey  gengærþ  hafua  tha  taki  alt
shall not provision have then take all

‘If a priest dies in the year when the bishop has no provision, the heir takes all the goods (and none are left to the bishop)’.

The trigger here is twofold: on the one hand heirs inherit only after somebody who died (dō), on the other hand, only provided the person who died owned some-
thing would there be something to inherit, (*præster*, as opposed to e.g. *þrel* ‘thrall (slave)’, for whose death no legal provisions are made in the OS laws).

The connection between death and inheritance is common in the legal prose, but it is not always marked. Example (47) is special since a contrast is highlighted between a potential and a real heir (the bishop vs. the relative).

Direct anaphora may be marked with a demonstrative: that is, the original use of the demonstrative that initiates the grammaticalization process. In indirect anaphora, there is no variation between -IN and demonstratives. Cross-linguistically, this context does not allow demonstratives, only definite articles. Yet there is one other item in variation with the article—namely the possessive pronoun.

(48) Beside the barn there is a little cottage. *The/Its roof* is leaking.

(Fraurud 2000b:246)

The variation in different marking of indirect anaphora can be studied in connection with a relatively common noun *arvi* ‘heir’, which often appears in indirect anaphoric contexts (when somebody’s death has been mentioned previously) in legal texts.

Table 15. The marking of the noun *arvi* / *arvinge* ‘heir’ in period I, legal texts

<table>
<thead>
<tr>
<th>Text</th>
<th>ø</th>
<th>adj</th>
<th>poss</th>
<th>-IN</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVL</td>
<td>23</td>
<td></td>
<td>12</td>
<td>-IN</td>
<td>arwumn</td>
</tr>
<tr>
<td>YVL (incl. ÆB)</td>
<td>13</td>
<td>7</td>
<td>20</td>
<td>9</td>
<td>1 (þriþia arfue ‘third heir’)</td>
</tr>
<tr>
<td>OgL (incl. ÆB)</td>
<td>1</td>
<td>12</td>
<td>26</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>DL</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 15 illustrates the use of different types of marking by the noun *arvi* (or *arvinge*) ‘heir’ occurring in indirect anaphoric contexts. To begin with, there is strong competition between bare noun and possessive in this context (1/3 of all instances in AVL includes a possessive). The percentage of unmarked forms is lower in the younger texts: in YVL there is more of a balance with a significant number of marked nouns. In OgL the possessive dominates over both the unmarked and marked forms.

52 Since ÆB *ärvdabalken* ‘the inheritance codex’ was not part of the text chosen for the study (see Chapter I, sources), and it has the highest frequency of the noun, it was added to the fragments searched for the purpose of this detailed study.

53 As for YVL.
OgL appears to be the most modern text in this regard. The only unmarked instance is in fact a negative polarity context (see Chapter 2):

(49) Nu dör man þæn sum egh hauær arua
now dies man this which not has heir

inrikis
domestically

‘If a man dies without an heir in the country’.

In such context the marking available is by a possessive, as in (50):

(50) Nu æn man ær utlænskær ok uet man
now if man is outlandish and know man

hans arua þa skal aruit standa þær til
his heir then shall legacy-DEF stand there to

ær hin kombær Nu æn man uet egh hans
if DET come now if man knows not his

harga
heir

‘If a man is a foreigner and his heir is known, the legacy shall be kept until the heir comes. If his heir is not known (..)’.

Interestingly enough, it seems as though modifying adjectives could be used instead of other marking such as -IN or a possessive. These are either reettir ‘right’ or sandir ‘true’, and can be found in all legal texts as modifiers of arvi. Examples (51) and (52) below are from OgL and DL.

(51) æn drapari kumi aldrigh i friþ fören rættær
but killer come never in peace before rightful

harga biþær firi hanum
heir offers for him

‘A killer is never at peace unless the rightful heir (of the victim) pleads for him’.

(52) Dör man wtlænskær af andru kunungs riki ær
Dies man foreign of other kingdom is

æi ærwingi æptir þa skal arf standa til nat
not heir after then shall legacy stand to night
'If a foreigner from another kingdom dies and there is no heir, his legacy shall remain untouched for one year and a night (from the day of his death). If the rightful heir comes with documents to prove his right to inherit, he shall take the inheritance'.

The possessive modifier may be either a pronoun or a noun (in the genitive), as in (53).

(53) Nu dör bonde (..) þa böte præstrin þrea
now dies peasant then pay priest-DEF three

markær arua bondans
marks heir peasant-DEF-GEN

‘If a peasant dies (..) then the priest shall pay three marks in fine to the peasant’s heir’.

The results of this brief study of the noun arvi ‘heir’ also show that the indirect anaphora is a context where -IN is in competition, not with the demonstrative as was the case in direct anaphora, but with the possessive (compare Chapter 2). It is to be noted, however, that bare nouns arvi or arvinge are also frequently employed in this context.

(54) Uærþær maþer dræpin ok af daghum takin þa skal
is man killed and of day taken then shall

uighi a þingi lysæ oc frafal
manslaughter on ting announce and loss

aruiningæ sigiæ
heir say

‘If a man is killed than the fact shall be made known on ting and to the heir’.

Typical examples of indirect anaphora involve whole-part relation between the first and the second referent or another, usually stereotypical, connection between the two. The relation between the verb ‘to die’ and the concepts of ‘heir’
and ‘legacy’ (the latter always unmarked) is not a whole-part one. It is an instance of a stereotypical connection between the verb and the nominal. Other examples of such stereotypical connections found in period I texts are collected in table 16.

Table 16. Indirect anaphora in period I, some trigger types

<table>
<thead>
<tr>
<th>Trigger type</th>
<th>Trigger</th>
<th>Translation</th>
<th>defNP</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>sak</td>
<td>charge</td>
<td>saksökande</td>
<td>plaintiff</td>
</tr>
<tr>
<td></td>
<td>kirkkyu</td>
<td>church</td>
<td>prestin</td>
<td>priest</td>
</tr>
<tr>
<td></td>
<td>tjuven</td>
<td>the thief</td>
<td>stölden</td>
<td>theft</td>
</tr>
<tr>
<td></td>
<td>tjuvnad</td>
<td>larceny</td>
<td>malseghande</td>
<td>plaintiff</td>
</tr>
<tr>
<td></td>
<td>sun manz</td>
<td>man’s son</td>
<td>fadheren</td>
<td>father</td>
</tr>
<tr>
<td></td>
<td>fe manz</td>
<td>man’s cattle</td>
<td>eghanden</td>
<td>owner</td>
</tr>
<tr>
<td></td>
<td>dom</td>
<td>sentence</td>
<td>domare</td>
<td>judge</td>
</tr>
<tr>
<td>Verbal</td>
<td>dör bonde (prest, landbo) /far bana af</td>
<td>dies yeoman (priest) / dies thereof</td>
<td>arfuanom</td>
<td>heir</td>
</tr>
<tr>
<td></td>
<td>madhur drepin</td>
<td>man killed</td>
<td>gerningen</td>
<td>deed</td>
</tr>
<tr>
<td></td>
<td>skuld krava</td>
<td>demand debt</td>
<td>guldit</td>
<td>gold</td>
</tr>
<tr>
<td></td>
<td>brennir man mans hus/ hugger hus</td>
<td>if a man burns another’s house / destroys another’s house</td>
<td>skada</td>
<td>damage</td>
</tr>
<tr>
<td></td>
<td>fördes död</td>
<td>the dead is brought</td>
<td>baren, liket</td>
<td>stretcher, corpse</td>
</tr>
<tr>
<td></td>
<td>wighia</td>
<td>marry, ordain</td>
<td>vixlenne</td>
<td>marriage, ordination</td>
</tr>
<tr>
<td></td>
<td>rida</td>
<td>ride</td>
<td>hästen</td>
<td>horse</td>
</tr>
<tr>
<td></td>
<td>henger</td>
<td>hang</td>
<td>galghan</td>
<td>gallows</td>
</tr>
</tbody>
</table>

A typical example of indirect anaphora was given in (48). There, the trigger was ‘cottage’ of which whole ‘the roof’ is a part. Triggers in indirect anaphora in period II are collected in table 17.
<table>
<thead>
<tr>
<th>Trigger type</th>
<th>Trigger</th>
<th>Translation</th>
<th>defNP</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nominal</strong></td>
<td>konungin</td>
<td>the king</td>
<td>drötningin (KM 251)</td>
<td>the queen</td>
</tr>
<tr>
<td></td>
<td>kirkio</td>
<td>church</td>
<td>til Patriarchans (KM 251)</td>
<td>to the</td>
</tr>
<tr>
<td></td>
<td>kejsaren</td>
<td>the emperor</td>
<td>drötningen (KM 254) / portalte (KM 255) / harith, skäggith (KM 255)</td>
<td>the queen / the doorway / the hair, the beard</td>
</tr>
<tr>
<td></td>
<td>all the swärdh</td>
<td>all the swords</td>
<td>oddana (KM 256)</td>
<td>the edges (blades)</td>
</tr>
<tr>
<td></td>
<td>spiwth</td>
<td>javelin</td>
<td>iärnith (KM 257)</td>
<td>the iron</td>
</tr>
<tr>
<td></td>
<td>anne</td>
<td>the river</td>
<td>watnith (KM 262)</td>
<td>the water</td>
</tr>
<tr>
<td></td>
<td>swärdh</td>
<td>sword</td>
<td>oddin, hiältith (KM 268)</td>
<td>the blade, the</td>
</tr>
<tr>
<td></td>
<td>Roland</td>
<td>Roland</td>
<td>blodhin (KM 279) / swärdhith (KM 287)</td>
<td>the blood / the sword</td>
</tr>
<tr>
<td></td>
<td>swärdhith</td>
<td>the sword</td>
<td>hiältina, brandin (KM 287)</td>
<td>the scabbard</td>
</tr>
<tr>
<td></td>
<td>et frwo kloster</td>
<td>a nunnery</td>
<td>altarith (KM 288)</td>
<td>the altar</td>
</tr>
<tr>
<td></td>
<td>smipir</td>
<td>a smith</td>
<td>äsionne (HML 190)</td>
<td>the hearth (forge)</td>
</tr>
<tr>
<td></td>
<td>stad</td>
<td>a town</td>
<td>gatunne, folkit (HML 191, 192)</td>
<td>the street, the people</td>
</tr>
<tr>
<td></td>
<td>eldin</td>
<td>the fire</td>
<td>lughanom (HML 192)</td>
<td>the flame</td>
</tr>
<tr>
<td></td>
<td>kättil</td>
<td>a cauldron</td>
<td>vatnith (HML 192)</td>
<td>the water</td>
</tr>
<tr>
<td></td>
<td>faghra käldo</td>
<td>beautiful water spring</td>
<td>vatneno (HML 193)</td>
<td>the water</td>
</tr>
<tr>
<td></td>
<td>hws</td>
<td>house</td>
<td>dörena (HML 222)</td>
<td>the doors</td>
</tr>
<tr>
<td><strong>Pronominal</strong></td>
<td>En ther höra skulle thera athäffwer</td>
<td>One which should listen to what they were doing</td>
<td>niwsnarin (KM 255)</td>
<td>the spy</td>
</tr>
<tr>
<td><strong>Verbal</strong></td>
<td>hugga</td>
<td>stab</td>
<td>swärdhith (KM 255)</td>
<td>the sword</td>
</tr>
<tr>
<td></td>
<td>Ridhu the thäðhan bort stak</td>
<td>they rode thence stabbed</td>
<td>ledhena (KM 264)</td>
<td>the way</td>
</tr>
<tr>
<td></td>
<td>Tha sprak henna hiärta sunder aff harm</td>
<td>then her heart burt from sorrow</td>
<td>oddin (KM 280)</td>
<td>the blade</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>likith (KM 288)</td>
<td>the corpse</td>
</tr>
</tbody>
</table>
As can be seen from the material presented, most of the indirect anaphora in period II are triggered by nominals and can be classified as either strict whole-part relation (swärdh ‘sword’—hiältena ‘the scabbard’) or stereotypical relation (kätil / anne ‘cauldron / river’—vatnit ‘the water’). There are some instances of indirect anaphora triggered by VPs. In period I, the majority of the examples include a stereotypical connection. It seems therefore that the indirect anaphora develops -IN marking in an opposite manner to the direct anaphora—it appears first in non-ambivalent contexts, and later to resolve doubts and link together parts of text; -IN is first used when the connection is easily accessible (more stereotypical). The spread of -IN leads from more to less stereotypical connections and in period II it is practically obligatory in textual contexts, including the indirect anaphoric.

(55) Biscopir huxar wighia en klærk til præst (. . .) han faar
bishop wants ordain one clerk to priest he gets
enkte aff vixlenne
nothing of ordination-DEF

(56) j samu stund førbes døb vt af stabenom (. . .)
in same moment led dead out of town-DEF
þa bøþ iohannes satia nípar barena
then bade Johannes set down bier/stretcher
ok løsa liket
and loosen corpse-DEF

(57) þa sanctus iohanes disputeraþe þesse skæl þa
when saint Johannes discussed these motives then
førbes en døbar tel graf þrætighi daghum før giptar
led one dead to grave thirty days before married
moþeren ok ænkian vini ok fraendar
mother-DEF and widow-DEF friends and relatives
fiollo for kne iohannis.
fell for knee Johannes-GEN
‘As St. John reasoned thus, a dead man was carried to the grave, who had been married thirty days earlier. His mother and widow, friends and relatives fell to John’s feet’.

In example (57) -IN is found with singular NPs and not with plural ones although the context for both is identical. It seems that in OS the usage of -IN is already limited to complete sets—while there is a single mother and a single widow, all friends and relatives would have to appear at the funeral for -IN to be used, since its usage presupposes completeness of the set (Hawkins 1978:157ff.).

Consider first the following data.

Bring the wickets in after the game of cricket.

Would I be satisfied if the hearer brought me only four or five of the six wickets? I would not. The wickets refers to all six.

I must ask you to move the sand from my gateway.

Would I be satisfied if only some were moved? As before, I would not. If my hearer only moved part of the sand away I would be justified in complaining: I thought I asked you to…

(Hawkins 1978:159)

In period III all textual functions are obligatorily marked with -IN (or a demonstrative or a possessive pronoun, see 4.5.1), as in (58-60), which are simple examples of stereotypical connection between the first and second referent.

(58) Riddarin thänkte mz sik at han war ey
knight-DEF thought with REFL that he was not
tha rätskr iptadhir ok gik atir til skriptafad heren
then well-confessed and went back to confessor-DEF
ok scriptadhe sik fulkomlika.
and confessed REFL fully

‘The knight thought he was not properly confessed and went back to his confessor and confessed fully’.

(59) Naar tässen hälgaiomfrwn war cristnat, oc aff
when this holy virgin-DEF was baptized and of
wandh myölkkenne aff modherliko bryste
turned milk-DEF of motherly breast

‘When this holy virgin was baptized and taken away from the milk from her mother’s breast’.
And then he drew the sword and slashed the one who carried the banner on his left shoulder, and cleaved (split) him so that the sword stood in the saddle’.

4.4.3. **Excursus: inalienable possession**

As we have noted previously, the most typical example of indirect anaphora is a whole-part relationship, e.g. a *house—the roof*. In this type of relationship are included instances of inalienable possession, involving body parts, clothes worn, weapons (there is a great cross-linguistic variation as to what elements may enter the inalienable possession relation), as well as kinship terms (see Chapter 2). Inalienable possession will be treated here separately, although it is a subtype of indirect anaphora, because it exhibits a different behaviour in OS. Even in MS there is a certain amount of variation between the definite and the possessive in the indirect anaphoric context. In MS the inalienable is usually expressed by the definite form of the noun but a reflexive possessive can also be used, usually within the same clause (see Chapter 1.5). In OS, however, if the noun is marked in this context at all, it is done exclusively by means of a possessive pronoun. Consider first some examples from legal prose in period I.

Then shall take an untame cow and lead it on a hill and shave all the hair off its tail⁵⁴.

---

⁵⁴ MS would in this sentence have *rumpan (svansen)* with a definite form, possessive *dess* ‘its’ is also possible.
Marked nouns are not found with human referents in legal prose:

(62) Værdær mahær huggin i houod ma seiae hinnu
is man hit in head may see membrane
kallæ bāji uapa bōtæ firi markum III
say both accident pay for marks three

(62) Værdær mahær huggin i houod ma seiae hinnu
is man hit in head may see membrane
kallæ bāji uapa bōtæ firi markum III
say both accident pay for marks three

‘If a man is hit in the head so that the membrane is visible, both (the attacker and the victim) say it was an accident, he shall pay three marks’.

Since there are few -IN forms in laws and the form is missing in many contexts (see 4.3 and 4.4) it may not be so surprising. However, even in the younger and more modern religious prose body parts and qualities are often represented as bare nouns, e.g.

(63) han gaf þøm ængen anzsurvey tan la þre dagha
he gave them no answer but lay three days
dumbe fiarpa daghen fik han atar mal
dumb fourth day-DEF got he back voice

(Bur 174)

‘He gave them no answer only lay dumb for three days, on the fourth he recovered his voice’.

In fact, the only text where the body parts are more or less regularly marked is Pent, which again shows the more advanced spread of -IN compared to other period I texts (as was the case with direct anaphora).

(64) Jacob swara thz monde wara gudz wili at thz
Jacob answers this may be god’s will that this
skulde skøt honom møta som han wilde Ysaac iæwar
should quickly him meet as he wanted Isaac hesitates
æn før malsins skuld oc bidhir han ganga
still for voice-GEN-DEF sake and asks him go
fram til sin
forward to self-GEN

(Pent 164)

‘Jacob answers that it may have been God’s will that this should happen so quickly as he wished. Isaac still hesitates because of the voice (he is uncertain whether it is Esau or Jacob) and asks him to come closer’.

55 MS: målet ‘the speech’, possessive less acceptable.
Hon  bidhir  sin  hærra  drikka  oc  førir  skøt  nidhir  
she  offers  her  lord  drink  and  brings  quickly  down  

carit  aff  axlomen  
vat  of  shoulder-DEF-DAT  and  runs  back  

‘She offers her lord a drink and brings quickly down the vat off her shoulders and 
rubs back to the well’.

However, we also find possessives in inalienable possession contexts in Pent.

thy  tordhe  ey  iacob  tagha  sik  hærbærge  aat  
thus  dared  not  Jacob  take  self  shelter  at  

quelde  vtan  lagdis  vt  a  markena  sidhan  
evening  but  laid  out  on  ground-DEF  after  

sol  war  sat  oc  lagde  en  sten  vndir  sith  
sun  was  sat  and  laid  EN  stone  under  his  

‘Thus Jacob did not dare find a shelter in the evening but laid down on the ground  
after the Sun had gone down and placed a stone under his head’.

In period II inalienable possession is almost exclusively marked by the possi-
ssive pronouns rather than the definite article; so is the case in KM. Consider  
the following examples:

Aff  the  klädhe  som  war  herra  hafðhe  om  sit  
of  this  cloth  which  our  lord  had  around  his  

hofwdh  tha  iwdha  korsfesto  han  
head  when  jews  crucified  him  (KM 252)  

‘(He gave him some) of the cloth which our Lord had around his head when the  
Jews crucified him’.

Jak  hafwer  en  hat.  tha  iak  hafwer  han  a  
I  have  a  hat  when  I  have  him  on  

56 In MS the definite would be preferable (om huvudet ‘around head-DEF’).
The development of textual functions


The only instance where a definite is used is (69), with the definite forms harith ‘the hair’ and skäggith ‘the beard’ denoting ‘his (=the emperor’s) hair’ and ‘his beard’.

(69) oc komber keysaren wt tha skal iak bläsa and comes emperor-DEF out then shall I blow

aff honum alt harith oc skäggith oc of him all hair-DEF and beard-DEF and

swa hans klädhe oc skal keysaren sta ater so his clothes and shall emperor-DEF stand again

nakudher naked

‘And when the emperor comes out I shall blow off him all of his hair, beard and clothes so that he stands naked again (as when he was born)’.

Finally, (70) shows the use of a possessive in an indirect anaphorical context, which may also be included in the inalienable possession (if weapon is considered a vital part of any rider).

(70) Ok olifernes ridher ärlika mote honum oc hiog and Olifernes rides honestly against him and struck

til hans mz sit swärdh oc fälde han to he-GEN with his sword and fell he-ACC
dödhan a iordh. dead-ACC on ground

‘And Olifernes rode openly againt him and struck him with his sword and fell him to the ground, dead’.

However, in comparison with KM, where the inalienable possession with animate referents was expressed mainly by possessives, there are many instances in SVM with -IN rather than possessives (e.g. mista maalith ‘lose one’s voice’, mista synen ‘lose one’s sight’). In fact, the variation in SVM is between -IN and possessive in inalienable contexts rather than possessive and bare noun.
In the youngest text in period II, Bo, we do not find a similar variation between the definite and possessive in inalienable possession contexts such as body parts, which are instead invariably marked with -IN.

The examples classified here as indirect anaphora are mainly inalienable possession with body parts and two instances of the noun altaret ‘the altar’ in connection with offering a sacrifice in a temple.

Inalienable possession in period III is still a context where either -IN or a possessive are used (74), possibly with greater preference for the definite (75). Bare nouns are no longer found in this context.

(71) oc fadheren tog han i hans hand
and father-DEF took him in his hand (SVM 115)

‘And his father took him by the hand’.

(72) The togh hon kesarens son i
then took she emperor-DEF-GEN son in

handena
hand-DEF-ACC (SVM 116)

‘She then took the emperor’s son by the hand’.

(73) Ok räkte piltin ihesus händrina aat
and reached boy-DEF Jesus hand-ACC-DEF to

foghomen ypystande öghonen til himilin
birds-DAT-DEF uplifting eyes-ACC-DEF to heaven-DEF

(Bo 9)

‘And the boy Jesus reached out his hand towards the birds, lifting his eyes to heaven’.

The examples classified here as indirect anaphora are mainly inalienable possession with body parts and two instances of the noun altaret ‘the altar’ in connection with offering a sacrifice in a temple.

Inalienable possession in period III is still a context where either -IN or a possessive are used (74), possibly with greater preference for the definite (75). Bare nouns are no longer found in this context.

(74) hwilkin ther wärra kan skall mista huffuodit (..) tha
which they defend can shall lose head-DEF then

skulde weland mistä sit huffuod
should Weland lose his head (Did 46)

‘Which they can defend shall lose the head (..) then should Weland lose his head’.

(75) tha tywffuen wträkte höghro armen äpther
when thief-DEF out-reached right arm-DEF after
The development of non-textual functions

We may conclude by saying that the use of -IN in indirect anaphora is not obligatory in period I and is mainly found by referents introduced not by another nominal, particularly not one stereotypically linked to the second referent, but mainly in connection with a VP. Neither is it established in inalienable possession, where the main variation is between possessives and bare nouns. This changes in period II when -IN starts appearing in the context and in period III all inalienable possession is obligatorily marked by either -IN or a possessive pronoun.

4.5. The development of non-textual functions

We have seen in 4.4 how -IN becomes obligatory in textual functions, both direct and indirect anaphora. Though it is found in these functions as early as period I, it is first in period III that it becomes obligatory. Though it is found in non-textual functions as well, the marking of uniques lags behind textual definites, and generics are only sporadically marked as late as in period III.

We use the term non-textual here to refer to all uses of -IN which lack a textual introduction, direct or indirect. This is not to say they are entirely text-independent, as will become evident in the discussion of unique referents in the religious prose.

4.5.1. Uniques

There are few instances of unique referents to be found in legal prose, which is only to be expected from their character. Those we find are usually unmarked, e.g. iord ‘mark, earth’ in contexts that in MS have the definite form.

(76) þa skal hana ur iorþ up taka ok lik then shall her from earth up take and corpse

hænna up skæra her up cut

‘She shall than be taken up from the earth (exhumated) and her body cut open’.
Apart from the specific meaning of 'the earth' (that was bought or sold or inherited etc., the noun iorþ was found twice in its marked form in OgL (whole text), both occurrences in what seems to be a lexicalized phrase, roughly corresponding to 'all the earthly goods'.

(77) þe haua allu þy foruærkat sum þe aghu
they have all this lost which they owned

ouan a iorþinne
over on earth-DEF

'(They shall then lose all their earthly possessions'.

Unique referents are more abundant in narratives, mainly in the religious prose. In Bur there are ample instances of sol ‘sun’, iorþ ‘earth’ and weruld ‘world’ alongside more specifically religious uniques like paradis ‘paradise’, natur ‘nature’, himinn ‘heaven’, dievul ‘devil’ or helviti ‘hell’. The distribution of -IN and bare nouns is given in Table 18.

Table 18. Distribution of -IN and -ø with unique referents in Bur

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>Bare noun</th>
<th>-IN</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>veruld ‘world’</td>
<td>30 (27 PP)</td>
<td>15</td>
<td>1 (de ‘this’)</td>
</tr>
<tr>
<td>sol ‘sun’</td>
<td>10 (comparison)</td>
<td>7</td>
<td>1 (pl), 1 (numeral), 3 (compounds)</td>
</tr>
<tr>
<td>iorþ ‘earth’</td>
<td>9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>natur ‘nature’</td>
<td>6</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>paradis ‘paradise’</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

The only instance where ‘world’ lacks -IN is after all.

(78) ena lund sannas iuþans lygh al væruld ær af
one place verify jewish lie all world is of

þinom fæþar bygd
your father built

(Bur 78)

Otherwise ‘world’ is definite, also after all:

57 The other occurrence is allu hauær hon foruærkat þy hon a ouan a iorþinne ‘all has she lost which she has on earth’ (OgL KE:26)
The development of non-textual functions

(79) ok þyngre var þu æn al þælデン
and heavier were you then all world-DEF

‘And heavier were you than all the world’.

There is more variation in the form of sol ‘sun’, but a closer inspection of the instances reveals that those without -IN appear solely in comparison contexts, like in example (80):

(80) Guz ængel kom een dagh tel โรงแรม lius som
 gods angel came one day to  Joachim light as
 sol ok  þroste  han  ræddan
 sun and comforted him scared

‘God’s angel came one day to Joachim, bright as sun, and comforted the scared man’.

The legal prose, again, prefers the unmarked form of the unique.

(81) þa  sum  sool  gangir  vndi  wid  wm  lõghardagh
then as sun goes under by on Saturday

þa  ær  inni  synnudags  hælg.
then is within Sunday-GEN holiday

‘Then, as the sun goes down on Saturday, the Sunday holiday begins’.

There are double as many occurrences of bare noun iorþ in Bur as of the noun with suffix. In this case, however, factors other than semantic ones seem to play a part—of the 30 bare nouns, 27 appear in PPs, as in the example below:

(82) cristoforus satte sin staf  i  iorþ
 Cristofor placed his stave in earth

It has been noted in earlier studies of definiteness (e.g. Larm 1936) that the use of -IN after preposition is of a later date. This seems to be confirmed by these results.

Terms that can be uniquely understood within the religious context, like ‘paradise’, ‘nature’ (= God’s creation), etc. show some degree of variation in marking but appear as bare nouns most often. This concerns terms that are loanwords in particular: the word paradis ‘paradise’ is always a bare noun in Bur.
They led me to the paradise and showed me the most glorious palace an eye may see.

It is to be noted that the word *paradis* is by no means a recent loan in Bur, since it is found in runic inscriptions as well. It may admittedly be difficult to fit in the inflectional paradigms.

Larm, who has studied the classical OS (= period I in this study) definites in greater detail, gives the following statistics for the unique referents in legal prose (table 19).

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>UL</th>
<th>OgL</th>
</tr>
</thead>
<tbody>
<tr>
<td>kunugrin ‘the king’</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>landit ‘the land’</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>biskuprin ‘the bishop’</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>præstrin ‘the priest’</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>domarin ‘the judge’</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>hundarit (hæræbit) ‘the district’</td>
<td>10</td>
<td>15</td>
</tr>
</tbody>
</table>

The NPs are organized from applying least to most ‘locally’. While ‘the king’ applies uniquely for everybody in a given country, ‘the district’ is only unique to its inhabitants. Larm’s results are difficult to judge because he does not state how many instances of a given word are found in UL and OgL and therefore what percentage of these are marked and what percentage are bare nouns.

Table 20\(^{58}\) shows the results of a similar study, covering three legal texts, UL and OgL studied by Larm and DL, all of approximately the same age (ca 1280). *Upplandslagen* (1297) is a relatively conservative legal text, whereas the contemporary *Östgötalagen* (1280) is much more modern linguistically.

\(^{58}\) For the sake of comparability, I follow Larm’s choice of referents for this study. It should however be noted that we could expect, and find, some of them, ‘priest’ and ‘judge’ in particular, in purely textual contexts as well, either in direct or in indirect anaphora. This is less probably for ‘king’ or ‘bishop’. Also, the noun ‘district’ is the only inanimate among the referents, which may influence the results.
DL (1320) is the youngest and yet linguistically, it occupies a medial position between UL and OgL.

Table 20. Unique referents in UL, OgL and DL

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>UL -IN</th>
<th>UL bare noun</th>
<th>UL total</th>
<th>OgL -IN</th>
<th>OgL bare noun</th>
<th>OgL total</th>
<th>DL -IN</th>
<th>DL bare noun</th>
<th>DL total</th>
</tr>
</thead>
<tbody>
<tr>
<td>kunugrin ‘the king’</td>
<td>0</td>
<td>115 (54 gen)</td>
<td>115</td>
<td>10</td>
<td>189 (98 gen)</td>
<td>199</td>
<td>2</td>
<td>82</td>
<td>84</td>
</tr>
<tr>
<td>landit ‘the land’</td>
<td>0</td>
<td>48</td>
<td>48</td>
<td>8</td>
<td>25</td>
<td>33</td>
<td>0</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>biskuprin ‘the bishop’</td>
<td>3</td>
<td>92</td>
<td>95</td>
<td>2</td>
<td>77</td>
<td>79</td>
<td>0</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>prestrin ‘the priest’</td>
<td>4</td>
<td>114</td>
<td>118</td>
<td>9</td>
<td>77</td>
<td>86</td>
<td>2</td>
<td>27</td>
<td>29</td>
</tr>
<tr>
<td>domaren ‘the judge’</td>
<td>6</td>
<td>33</td>
<td>39</td>
<td>9</td>
<td>0</td>
<td>9</td>
<td>8</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>hundarit (hæraþit) ‘the district’</td>
<td>9</td>
<td>87 (41 gen)</td>
<td>96</td>
<td>15</td>
<td>68</td>
<td>83</td>
<td>12</td>
<td>4 (gen only)</td>
<td>16</td>
</tr>
</tbody>
</table>

As can be seen in table 20, the noun konung ‘king’ is very seldom marked with -IN. One particular context for this noun is the coordinated NP, e.g. ‘king and bishop and district’, stating who is to receive fines for crimes committed by the one found guilty. This context is highly lexicalized, since the fines are traditionally divided into three parts, one part to be paid to the king, the other to the bishop and the third to ‘all men’ (for public benefit) or to the plaintiff. There are some minor variations, the fines being paid to district rather than the bishop in lay matters (non-religious). But the type is basically the same. There is also some variation in the grammatical form which may be dative, as in ‘to the king etc.’ or nominative, as in ‘the king has/takes fines.’

If a highly lexicalized context shows new forms, this may be considered a sign of their advancement (in terms of grammaticalization). This particular context involves referents that may be identifiable for different reasons (the plaintiff is either mentioned before or may be indirectly inferred from the legal case discussed earlier; the district would be ‘the closest one, the most relevant one, his’; king and bishop are representatives of higher power, uniquely identified in the whole country). We will now study these ‘fine payment-contexts’ in OS legal prose in more detail.

Both in AVL and YVL there are only sporadic instances of -IN, and most beneficiaries are unmarked.
Dræpær maþær man (..) bøti huarti konongi allær hæræþi

killed man man pay not king or district

‘If a man kills another (accidentally), he shall pay the fines neither to the king nor to the district’.

bótæ viþ hæræþ IX markær ok sva uþ konong

pay by district nine marks and so by king

‘(He shall) pay nine marks to the district and same amount to the king’.

If some referent is marked, it will usually be the most ‘local’ one, the plaintiff rather than the bishop, the parish rather than the king.

ær præster forfalla lös bøte uith biskup sak sina

is priest excuse loose pay by bishop charge his

oc sacsökianum thre marker and plaintiff-DEF three marks

‘If the priest has no excuse (for not giving the anointment to the dying member of his parish) he shall pay the bishop and the plaintiff three marks’.

Præster ma ingin v kirkyu uræka vtan þæn biskuper

priest may none out church throw without this bishop

hauer forbuthat. (..) Vræker annan dagh. bøte mark biskupi

has forbidden throws other day pay mark bishop

oc sua kononge oc sua hærathe.

and so king and so district

Uráeker thrithia dagh. bøte nio mark soknenne.

throw third day pay nine mark parish

oc IX kononge. ok IX hærathe

and nine king and nine district

‘A priest may not throw any man out of the church, except one that has been forbidden (to enter) by the bishop. (..) If the priest throws the man on the second day, he shall pay one mark to the bishop and likewise to the king and the district. If he throws him out on the third day (for the third time), he shall pay nine marks to the parish and nine to the king and nine to the district’.

We have noted before that OgL is considered the most advanced linguistically among OS legal texts. With respect to the ‘fine payment-phrase’ it shows great variation, as examples (88a) and (88b) show.
The development of non-textual functions

(88a) ok böta malsæghanda ok hæraþe ok kununge ok
and pay plaintiff and district and king and

biskupe
bishop

(88b) ok þerra gozs sum þe aghu skal skiptas
and their goods which they own shall divide-PASS

i þrea lyti en lutin malsæghandanum
in three parts one plaintiff-DEF

annan kununginum þripia hæræþinu
other king-DEF third district-DEF

‘And (he shall) pay the plaintiff, the district, the king and the bishop’.

‘And the goods that they own shall be divided into three parts: one for the plaintiff, the other for the king and the third for the district’. 59

DL, a text with relatively high frequency of -IN, on the other hand, does not display a single instance of konungin in this context though it marks others beneficiaries.

(89) J twæskipte taki halft kunungir oc halft
in two-division take half king and half

hundarit
district-DEF

‘When the fines are divided in two, the king takes half and the district the other half’.

The unique marking seems to spread from the referents that are more local (like the plaintiff, also an animate referent, or the district) to those less local (the king). Animacy does not seem to play a part here, as examples such as (90), with the only inanimate noun also being the only one marked, are highly frequent.

(90) taki en lut kunungir annan biscopir þripia
take one part king-NOM another bishop-NOM third

59 Note the marked form of luter ‘part’ in the phrase en lutin ‘one part-DEF’. This type of marking will be discussed in 5.9.1.
The king takes one part, the bishop the other and the district the third.

We must assume a relatively high level of lexicalization of the phrases, since examples like (91) are also found.

(91) giui kunungenum þæt han toki oc börí XL
give king-DAT-DEF this he took and pay 40

markir kununge
marks king-DAT (DL Þing 11)

(he shall) give the king what he took and pay a fine of 40 mark to the king.

Apart from ‘the sun’, Old Swedish unique concepts also include hafweth ‘the sea’, oknen ‘the desert’ and occasionally even skogen ‘the forest’.

(92) Langt ær til kirkiu fara guzziuiu sculu barn wacta.
long is to church travel caretaker should child mind

Sæghia swa guzziuiu. wi ærum æi för máeþ
says so caretaker we are not for with

þæssu barnni ywi scoghin
this child over forest-DEF (DL KB:6)

‘If it is a long way to church, the caretaker shall take care of the child. If the caretaker says: we cannot travel across the woods with this child’.

In period II they are found in both religious prose and the narratives.

(93) Sidhan skyndadhe the sina färðh thadhán oc
then hurried they their journey thence and

komo til marsiliam. oc komo alle hilbrighdho
came to Marseille and came all healthy

ower hafwith oc komo til ierusalem oc
over sea-DEF.ACC and came to Jerusalem and

toko sik ther härberghe
took themselves there shelter (KM 251)
‘They then hurried from there and came to Marseille and came in good health over the sea and to Jerusalem where they took shelter’.

At a first glance, (93) could be classified as indirect anaphora, since the use of the definite on the noun *haf* ‘sea’ seems to be triggered by the place name Marseille, a port, thither and whence journeys are made by sea. There is a piece of information in the text that induces the definite. However, a comparison with other texts shows that the concept could in fact be unique to the people who wrote (and read) these stories and was simply to be understood as a great mass of sail-able water. In SVM the noun *haf* is consistently definite even though no port towns are mentioned and no other (linguistic) context given.

Even though unique referents in period II are commonly marked with -IN, some variation remains, as examples (94) and (95) from KM show.

(94) the waro swa onde men at iordhin opnadhis
they were so evil men that earth-DEF opened

fore them oc swenko swa nidher til häluitis
for them and sank so down to hell-GEN

(KM 270)

‘They were such evil men that the earth opened before them and they sank to hell’.

The usage is not yet fully obligatory, as the following can also be found in KM:

(95) Ok olifernes ridher ärlika mote honum oc hiog
and Olifernes rides honestly against him and struck

til hans mz sit swärðh oc fälde han
to he-GEN with his sword and fell he-ACC

dödhan a iordh.
dead-ACC on ground

(KM 270)

‘And Olifernes rode openly againt him and struck him with his sword and fell him to the ground, dead’.

It is possible though that (95) is a lexicalized phrase, since there are three more similar examples in KM, including the identical wording of ‘fell him to the ground, dead’.

A clearly unique referent, that is already definite in earlier texts, *soll* ‘sun’, appears in the same narrative (twice throughout the text\(^\text{60}\)) without the article:

\(^\text{60}\) There is a third occurrence in a comparative phrase *som en soll* ‘(bright) as a sun’ which does not refer to ‘the Sun’.
The unmarked form may also be due to the fact that this is a negative polarity context (after negation). We should note at this point that even if the negative polarity context influences the reading of the noun, the uniques in MS are not influenced by it.

We noted previously that the loanwords such as paradis do not take -IN as easily as indigenous ones. A closer look at Bo reveals this to be true—although the frequency of the suffix is very high and -IN seems obligatory with unique referents, paradis appears only once with the suffix and in the remaining (7) instances is a bare noun.

In period III we find uniques such as sol ‘sun’, wæruld ‘world’ and iord ‘earth’ always marked with -IN, including negative polarity contexts.
In Table 21, frequencies of the different types of marking of the unique referents are given. The words studied are the inherited *veruld*, *sol*, and the loans *natur* and *paradis*. They are studied in religious texts from all periods.

**Table 21. Distribution of -IN and -ø with unique referents in religious prose**, periods I-III

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>Period I</th>
<th>Period II</th>
<th>Period III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bare noun</td>
<td>-IN</td>
<td>Other</td>
</tr>
<tr>
<td><em>veruld</em></td>
<td>6</td>
<td>206</td>
<td>6</td>
</tr>
<tr>
<td><em>sol</em></td>
<td>46</td>
<td>81</td>
<td>–</td>
</tr>
<tr>
<td><em>natur</em></td>
<td>260</td>
<td>37</td>
<td>–</td>
</tr>
<tr>
<td><em>paradis</em></td>
<td>103</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

As can be seen from the results presented in Table 21, the noun ‘world’ is regularly marked with -IN as early as the earliest religious prose; some unmarked instances have been found (though none in period II). It is the only unique that can be marked other than by -IN—mainly by the demonstrative *denna* or possessive *min* ‘mine’. In these cases there is a contrast implied between the world of earthly sorrows and that of the glorious future in heaven, thus it cannot quite be treated as a unique referent.

The noun *sol*, on the other hand, although also marked in the earliest texts, is unmarked in over 30% of instances in period I and occasionally unmarked in period II (99) and even in period III (100).

(99) | at | han | atirkom | til | rääät | samuit | oc | gik |
|----|----|------|------|------|--------|    |    |
|    | he | returned to | right | conscience | and | went |

(100) | J | then | daghen | kwnnom | wi | enkte | swara | nar |
|    | in | this | day-DEF | can | we | nothing | answer | when |

---

61 Period I: Bur and Pent B (complete texts), period II: HML, Jart and Bo (complete texts), period III: Linc NR and Linc JB (complete texts).
‘On that day we can answer nothing when both heaven and earth, fire and water, sun and moon, day and night and everything in the whole world stand and witness against us and our sins’.

Of the two loanwords studied, *natur* ‘nature’ seems to be more easily incorporated into the inflectional system, taking -IN regularly in period II, whereas *paradis* ‘paradise’ remains unmarked throughout all periods and the only marked instance involves a strong deixis (*detta paradysit* in example (97)). Throughout the religious prose we find uses of -IN which are clearly non-textual and are yet difficult to classify as uniques. Consider the following period III example:

(101) oc främia sik i dygdhomen oc godhom and hone self in virtue-DAT-DEF and good
gärninggom deeds (Linc NR 277)

‘And to hone oneself in the virtues and good deeds’

Also references made to the commonly known biblical characters and events are always marked with -IN.

(102) Laudate dominum in sanctis eius, konnogxlike propheten dawid
Laudate dominum in sanctis eius, royal prophet-DEF David
i psalmenom sigher tässen ordhen
in psalms-DAT-DEF says these words-DEF (Linc NR 276)

‘Laudate dominum in sanctis eius, the royal prophet David says these words in the psalms’.

These atypical uniques can occasionally be found outside religious prose as well, e.g. in the narrative PK. However, they occur only in the introductory part
of the chronicle, presenting the world’s history (among others, the flood and the raising of the Babel tower) as known from the Bible, which are presented in PK as part of the history of Sweden, since the Swedes in that text are considered to be the descendants of Noe’s son, Japhet.

(103) Hwilka affödha och äpterkomande förgato och änglana
which descendants and progeny forgot and angels’

downfall och flodhen som gangen war öffwer
and flood-DEF which gone was over

alt mankönit och begynthe högfärdhinna
all mankind and began vanity-DEF torn

oppa ena mark som tha kalladis sennar och
upon a ground which then was-called Sennar and

nw kallas babilionia
now is-called Babylon

(PK 220)

‘These descendants forgot both the downfall of the angels and the flood which came on all the mankind and began (building) a tower of vanity in a place once called Sennar and today known as Babylon’.

This use of -IN will be discussed in more detail in 4.6.3.

4.5.2. Generics

Finding -IN used generically in period I is a difficult task. The use of nouns maðr ‘man’ and bonde ‘yeoman’ in legal prose can be said to be of a generic nature, since no specific ‘man’ is meant. However, Wessén points out that the use of the nouns in legal texts is highly specific and that they are not to be understood as generic uses. The noun bonde will be studied separately in 4.6.1. Suffice it to say that Wessén’s ideas seem to be confirmed here.

Otherwise, isolated uses of -IN nearing generics can be found as early as in Bur (period I).

(104) En af giri ok lusta likamans kumba {heþers}
but of greed and lust body-DEF-GEN comes honour

spiæl
ruin

(Bur 183)

‘But of the greed and lust of the body comes the corruption of honour’.
In period II, some -IN generics can be found in Bo. Note, however, the variation in the marking of generics in the short passage quoted in (106), where we find bare noun, -IN and EN with apparently the same function.

Apart from the generic uses of both -IN and en in the previous example, we find marked forms of nouns ‘purity’, ‘holiness’ and ‘virtue’. These belong to a wider category of nouns which are almost always definite in Bo and other religious texts, although they have not been introduced linguistically before, neither is the use generic. We take a closer look at the whole group in 4.6.3.

We have seen in Chapter 1.5 that there are five different ways of expressing generic reference in MS: by bare noun, definite or indefinite singular and definite or indefinite plural (see also Pettersson 1976). As late as in period III this complex system is yet to be formed. The variation in the marking of generics is largely between -IN and bare noun in the singular, with occasional uses of EN.
The development of non-textual functions

4.5.3. From textual to non-textual uses of -IN

We have previously seen that -IN appears in all textual contexts in the earliest texts, including indirect anaphora, but it is found in non-textual uses only in texts of later dates. The textual functions are therefore older than the non-textual ones (see also Chapter 2). The question now is what types of contexts allow the spread of -IN from textual to non-textual functions and whether these can be found in period I material.

In 4.5.1 we have discussed, at some length, the different types of unique referents to be found in legal texts, concerning more and less local authorities (the king, the bishop, the district). Among these the noun ‘priest’ belongs to the most local category, since its ‘uniqueness’ is limited to one parish. Bearing that in mind, consider the following example:

---

**4.5.3. From textual to non-textual uses of -IN**

We have previously seen that -IN appears in all textual contexts in the earliest texts, including indirect anaphora, but it is found in non-textual uses only in texts of later dates. The textual functions are therefore older than the non-textual ones (see also Chapter 2). The question now is what types of contexts allow the spread of -IN from textual to non-textual functions and whether these can be found in period I material.

In 4.5.1 we have discussed, at some length, the different types of unique referents to be found in legal texts, concerning more and less local authorities (the king, the bishop, the district). Among these the noun ‘priest’ belongs to the most local category, since its ‘uniqueness’ is limited to one parish. Bearing that in mind, consider the following example:
(110) Hua inne uill ligia i kirkyu gifui hænni þre who inside wants lie in church give her three
mark oc praestinum sijn læghorstath oc marks and priest-DAT-DEF his burial ground and
spøre ey lænspraestin at ask not vicar to (YVL KB:23)

‘Who wishes to be buried inside the church (and not in the churchyard) should pay three marks to the church and to the priest for the burial place, without consulting the vicar.’

In example (110) we see a context that has a twofold interpretation: on the one hand the definite form of ‘priest’ may be a matter of indirect anaphora—where there is a church, there is a priest, the priest of the church in question. On the other hand it may be ‘the closest priest’—local uniqueness.

That the more local uniques are earlier marked with -IN has already been discussed in 4.5.1. It is very clear from the example (111) below.

(111) Ey ma kirkyu iorth sæliæ ælla skipta vtan not may church mark sell or divide without
biskups loff oc praestins oc bishop-GEN permission and priest-GEN-DEF and
soknamanna villæ vtan þa sum gifs parishmen-GEN will without this which give-PASS
til kirkyu bygningh hana skulu soknamæn sæliæ to church maintenance her shall parishioners sell
æn kirkia thorf bygning uith if church needs building to (YVL KB:67)

‘The ground belonging to the church may not be sold or divided without the permission of the bishop and the priest and the parishioners unless the sale is to bring in the money for the maintenance of the church, then the parishioners may sell if the church is in need of maintenance’.

In this case, there is a difference between the priest and the bishop; the latter, although also unique and indirectly anaphoric (the bishop who oversees the particular church) is unmarked.

(112) Ganger præster i annars sokn oc bær i bok goes priest in other’s parish and carries in book
4.6. Anomalous uses of -IN

The aim of this section is to report the uses of -IN that are surprising from the MS perspective but even more so, that are difficult to account for within the model we have adopted so far—differentiating between textual and non-textual definiteness. Despite the adjective ‘anomalous’ in the headline regularity can be found in this discrepant use of -IN.

4.6.1. Bondi in legal prose

Let us first consider the use of the noun bonde ‘yeoman’ in one fragment (kirkiubalken, ‘the church codex’) of DL (legal text, period I).

(112) Bondi skal tiunda præsti sinum (..) Böndir skulu kirkiu
yeoman shall tithe priest his yeomen shall church

gewa (..)
give

Bondin scal reta tiund giora (..) Hwar bonde scal
yeoman-DEF shall right tithes make each yeoman shall

giora
make

‘A yeoman shall pay tithes to his priest (..) Yeomen shall give (money) to the church (..) The yeoman shall pay proper tithes (..) Each yeoman shall make (..)”.

In this fragment, there is great variation in expressing similar content, either by use of bare noun in the singular, bare noun in the plural, a marked noun or
a pronoun *hwar* ‘each’. No subtleties of legal prose can be attached to this variation in form and we are therefore left to surmise that the system is in a phase of change and the scribe of the text makes full use of the forms at his disposal.

The use of -IN in contexts such as (112) is particularly surprising. The example is not an aberration, however, and similar ones can be easily found.

(113)  Nu  kan  **bondin**  stulin  uarþa
now  can  yeoman-DEF  robbed  become  (ÖgL VM:32)

‘A yeoman may be robbed (=If a yeoman is robbed)’.

(114)  Nu  dör  husfru  **bondans**
now  dies  wife  yeoman-GEN-DEF-GEN  (ÖgL GB:17)

‘Now dies yeoman’s wife (=If a yeoman’s wife should die)”.

(115)  Nu  ær  þæt  **klokkarans**  skuld  bæra  bok
now  is  it  sacristan-DEF-GEN  sake  bear  book
ok  stol  i  sokn  meþ  præste
and  stool  in  parish  with  priest  (ÖgL KB:6)

‘It is the sacristan’s task to carry the Bible and the stool when the priest makes a round in the parish’.

(116)  Nu  talar  lænsman  til  **bondans**  om  bloþ
now  talks  sheriff  to  yeoman-GEN-DEF-GEN  about  blood
wecneng
spillage  (VmL KB:26)

‘If a sheriff speaks to a yeoman about blood spill’.

MS would, in corresponding instances, use the indefinite rather than the definite article or possibly a bare noun. This unusual—from the MS perspective—usage of -IN is limited to the legal prose and therein mainly to the weak noun *bonde* ‘yeoman’ (frequently used as an equivalent of ‘man’), particularly in the oblique cases.

This overuse is observed in Larm 1936:100f, who wishes to see it as means of marking the case of the noun and claims it to be limited mainly to the genitive case. This interpretation is based on the fact that *bonde* being a weak noun, its marked form is clearly either genitive (*bonda-ns*), dative (*bonda-nom*) or accusative (*bonda-n*) while the bare noun is ambiguous and may be interpreted as dative, genetive or accusative depending on the context (*bonda*, oblique). According to Larm such usage of -IN is sporadic.
A closer inspection of the legal prose however reveals that the use is fairly frequent in a number of legal texts, though not in all. Table 22 shows strikingly high percentages of the marked form of the noun. In DL as much as ca 75% of all instances of the lexeme are marked with -IN, almost 50% in OgL and ca 67% in UL.

Table 22. Distribution of the noun bondi with and without -IN in DL, UL and OgL.

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>DL -IN</th>
<th>DL bare noun</th>
<th>DL total</th>
<th>UL -IN</th>
<th>UL bare noun</th>
<th>UL total</th>
<th>OgL -IN</th>
<th>OgL bare noun</th>
<th>OgL total</th>
</tr>
</thead>
<tbody>
<tr>
<td>bonde ‘yeoman’ NOM</td>
<td>77</td>
<td>26</td>
<td>103</td>
<td>88</td>
<td>166</td>
<td>254</td>
<td>109</td>
<td>118</td>
<td>227</td>
</tr>
</tbody>
</table>

Larm explains the overuse of the marked form by the desire to make the case form explicit (though he only mentions the genitive). Bonde is a weak masculine noun, nd-stem, which in the singular only differentiates between the nominative (spelled bondi or bonde) and oblique (bonda) forms. On the other hand, there is a maximal contrast in the ‘definite’ paradigm, where 8 different clusters of functions are expressed by 8 different forms. For the unmarked noun, on the other hand, the contrast is between nominative and oblique (two forms for four functions) in the singular and between the nominative/accusative, the dative and the genitive (three forms for four functions) in the plural. In other words, the definite paradigm makes the case differences clear in both numbers, while the bare noun shows a high level of syncretism in the singular.

Table 23. The inflectional paradigm of bonde ‘yeoman’

<table>
<thead>
<tr>
<th>Number</th>
<th>Case</th>
<th>Unmarked</th>
<th>-IN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg</td>
<td>Nom</td>
<td>bondi (bonde)</td>
<td>bondi-n</td>
</tr>
<tr>
<td></td>
<td>Gen</td>
<td>bonda</td>
<td>bonda-ns</td>
</tr>
<tr>
<td></td>
<td>Dat</td>
<td>bonda</td>
<td>bonda-num</td>
</tr>
<tr>
<td></td>
<td>Acc</td>
<td>bonda</td>
<td>bonda-n</td>
</tr>
<tr>
<td>Pl</td>
<td>Nom</td>
<td>bonder</td>
<td>bonder-ni(r)</td>
</tr>
<tr>
<td></td>
<td>Gen</td>
<td>bonda</td>
<td>bonda-nna</td>
</tr>
<tr>
<td></td>
<td>Dat</td>
<td>bondum</td>
<td>bondum-in</td>
</tr>
<tr>
<td></td>
<td>acc</td>
<td>bonder</td>
<td>bonder-na</td>
</tr>
</tbody>
</table>
Table 24 reports the results of the search for the nominative bonde (in all spelling varieties) and the marked bondin (also spelled bonden). The study was made for all legal texts from period I including the first country-wide law, Magnus Erikssons landslag (ca 1350). Altogether 9 texts were searched in extenso.

Table 24. Noun bondi in the nominative in all OS legal texts.

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>AVL</th>
<th>YVL</th>
<th>DL</th>
<th>OgL</th>
<th>UL</th>
<th>HL</th>
<th>YVest</th>
<th>Sdml</th>
<th>MEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>bonde</td>
<td>35</td>
<td>66</td>
<td>16</td>
<td>90</td>
<td>99</td>
<td>9</td>
<td>62</td>
<td>59</td>
<td>71</td>
</tr>
<tr>
<td>bonde</td>
<td>1</td>
<td>13</td>
<td>32</td>
<td>62</td>
<td>46</td>
<td>–</td>
<td>17</td>
<td>10</td>
<td>23</td>
</tr>
<tr>
<td>Percentage of -IN forms</td>
<td>2,8%</td>
<td>16,5%</td>
<td>66,7%</td>
<td>40,8%</td>
<td>31,7%</td>
<td>21,5%</td>
<td>14,5%</td>
<td>24,5%</td>
<td></td>
</tr>
</tbody>
</table>

The results presented in table 24 show that a high percentage of the marked form is also found in the nominative. In DL, the most progressive text in this respect, the marked form clearly dominates the unmarked one.

The results of the search for bondi in oblique cases (genitive, dative, accusative) are reported in table 25.

Table 25. Noun bondi in oblique cases in all OS legal texts

<table>
<thead>
<tr>
<th>Case marking</th>
<th>Case</th>
<th>Form</th>
<th>AVL</th>
<th>YVL</th>
<th>DL</th>
<th>OgL</th>
<th>UL</th>
<th>HL</th>
<th>YVest</th>
<th>Sdml</th>
<th>MEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>unmarked</td>
<td>gen/dat/acc</td>
<td>bonda</td>
<td>26</td>
<td>21</td>
<td>7</td>
<td>27</td>
<td>62</td>
<td>56</td>
<td>23</td>
<td>47</td>
<td>21</td>
</tr>
<tr>
<td>marked -IN</td>
<td>gen</td>
<td>bondans^62</td>
<td>–</td>
<td>3</td>
<td>29</td>
<td>30</td>
<td>24</td>
<td>8</td>
<td>9</td>
<td>27</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>dat</td>
<td>bondan-nom</td>
<td>–</td>
<td>–</td>
<td>5</td>
<td>11</td>
<td>14</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>acc</td>
<td>bondan</td>
<td>–</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Percentage of -IN forms</td>
<td>0%</td>
<td>16%</td>
<td>83,3%</td>
<td>63,5%</td>
<td>39,8%</td>
<td>18,8%</td>
<td>39,5%</td>
<td>39%</td>
<td>66,7%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^62 YVL has one instance of bondens, probably misspelled. In DL there are three instances spelled bondands.
If we look at the percentage of all bonda and bondans/bondanom/bondan instances, we will note that the marked forms make up for between 16% and 83% (disregarding AVL). Four texts have particularly high frequency of -IN in the genitive: DL, OgL, HL and SdmL. For MEL, it is the dative that dominates in the marked uses. However, as we mentioned before, if the primary function of the -IN was to disambiguate the case form, we would not expect to find it in the nominative, which is the only unambiguous case form for bondi. And yet, the dominance of -IN in the nominative is overwhelming in DL and the percentage also very high in OgL, UL and MEL. In fact, the only text where -IN is not used is HL. We can further see that the over-representation of -IN is mainly found in DL and OgL, with lower numbers in YVestL (a continuation of DL!) and SdmL.

Table 26. Marked and unmarked forms of the noun klokkare ‘sacristan’ in OS legal texts

<table>
<thead>
<tr>
<th>Case marking</th>
<th>Case</th>
<th>Form</th>
<th>YVL</th>
<th>DL</th>
<th>OgL</th>
<th>UL</th>
<th>HL</th>
<th>YVest</th>
<th>SdmL</th>
<th>MEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>unmarked</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nominative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gen/dat/acc</td>
<td>bondare</td>
<td>2</td>
<td>10</td>
<td>8</td>
<td>16</td>
<td>4</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>klokkara</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>16</td>
<td>6</td>
<td>14</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>marked -IN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>acc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>klokkarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>klokkarans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>klockarenum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>klockaran</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
It does not seem, therefore, as if the primary function of -IN was to make the case form explicit. But to make sure we may take a look at other weak nouns of relatively high frequency, which would need some clarification: domare ‘judge’ and klokkare ‘sacristan’.

Table 27. Marked and unmarked forms of the noun domare ‘judge’ in OS legal texts

<table>
<thead>
<tr>
<th>Case marking</th>
<th>Case</th>
<th>Form</th>
<th>YVL</th>
<th>DL</th>
<th>OgL</th>
<th>UL</th>
<th>HL</th>
<th>YVest</th>
<th>Sdml</th>
<th>MEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>unmarked</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nominative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gen/dat/acc</td>
<td>domare</td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domare</td>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domare</td>
<td></td>
<td>19</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domare</td>
<td></td>
<td>12</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domare</td>
<td></td>
<td>7</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domare</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>marked -IN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nom</td>
<td>domarin</td>
<td></td>
<td>5</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gen</td>
<td>domarin</td>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dat</td>
<td>domarin</td>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>acc</td>
<td>domarin</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domarin</td>
<td></td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domarin</td>
<td></td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domarin</td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>domarin</td>
<td></td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since both nouns have lower frequency than bonde, only absolute values are given and no percentages. It is clear that the nouns klokkare and domare do not exhibit the same pattern as bonde and the use of bare nouns dominates. Admittedly, in OgL it only appears as domarin and in DL the -IN forms dominate over the unmarked ones, even in the nominative. These two texts prove again to be the most advanced in this particular use of -IN. In other texts the unmarked forms dominate.

The use of -IN for stating the case form clearly is too strong a hypothesis, since weak nouns have always been weak and the syncretism of forms alone has not led to the reformation of the system. Let us consider some more examples.

Definite bondin appears in many contexts where specific en would be employed in MS, e.g.

(117) Warþir bondin dræpin be yeoman-DEF killed (DL MK:3)

‘If a yeoman is killed’.

(118) Dör bondin oc barn ær ængti æpti dies peasant-DEF and child is none after (DL GB:12)

‘If a peasant dies without an heir’.

The usage is not definite, but it may be regarded as specific. In MS it would be natural to use the indefinite article here. The usage is not limited to the noun
Anomalous uses of -IN —although this has the highest frequency—but can be found with other nouns as well, e.g. *nempdemaðr* ‘juryman’.

(119) Fylgher eig nempdæmaþrin sinnæ gengiærþ oc
follows not juryman-DEF his provision and
almenpnigs öre innan forstæ endaghom
common-land öre before first day-and-year
böte tua öræ bem gengiærþinæ a æller
pay two öræ him provision owns or
viti forfall sin meþ tolf mannæ eþe.
witness loss his with twelve men oath (YVL FB:42)

‘If a juryman does not fulfill his obligations and taxation in due time, he shall pay two öre fine where it is owned or witness the loss with the oath of the twelve’.

*Nempdemaðr* is a strong noun and in no need of clarification of the case form (there is no syncretism in the singular). It cannot be interpreted as indirect anaphora (= the closest) or unique because there were more than one juryman for each district, therefore the interpretation of -IN as an MS definite article is unsuccessful. A specific interpretation (comparable with the MS use of the indefinite article) is however possible.

If we recall the stages of grammaticalization of the definite article as presented in Greenberg 1978 (repeated as (120) below, see also Chapter 2), we see that the development goes through stage I (the definite article) to stage II (the specific marker).

(120) Stage 0 Stage I Stage II Stage III
demonstrative definite article specific article noun marker
(after Greenberg 1978, see also de Mulder and Carlier 2011:525)

However, it has also been noted that the grammaticalization from the definite article to the specific marker may be less likely to occur in languages that have also developed indefinite article (de Mulder and Carlier 2011:534). In other words, the presence of the indefinite blocks the further grammaticalization of

---

63 As we have noted in Chapter 1 and 2, the definite article implies completeness of reference, totality. Therefore it is correct to say

He bought a new house and had to relay the roof.
because there is one roof per house. It is however incorrect to say

He bought a new house and had to change the window.
because there are typically more than one window to choose from in a house.
the definite. It has further been noted that in the absence of the indefinite article, the development may proceed straight to the specific marker, and the form never becomes a definite article (in the sense of marking known/identifiable/unique information) but rather develops from a demonstrative to a specific marker (Himmelmann 2001).

Since there is no indefinite article in period I, not even in an embryonic form (as far as the sources tell us, see Chapter 5), -IN has, so to speak, more room to spread. It therefore seems to be at least partly used as a specific marker rather than a definite article.

The overuse of -IN may therefore be interpreted in the following ways:

- an attempt at clarification of syncretic case forms. Argument against: -IN found also in the nominative; no such regularity for other weak nouns such as klokkare and domare.
- a regional (dialectal) peculiarity. Argument against: the continuation of DL (representing the same regional variety) shows quite discrepant results, and the overuse is also found in MEL.
- an early stage of -IN grammaticalization, by-passing the definite article stage and instead going straight to specific marker. The development checked by the grammaticalization of en, -IN reinterpreted as ‘definite’ in opposition to en ‘indefinite’. Argument against: the overuse limited to a handful of nouns.

Possibly, all three explanations are partly correct and none can alone explain this discrepant use of -IN. Similarly unfamiliar use of -IN can also be found in Old Icelandic—in his OI grammar Heusler describes it as ‘opposite’ (umgekehrt) to the definite article in Modern German (Heusler 1950:126), giving some examples:

(121)  kuezk eiga dýret eitthuert
sveys-REFL own animal-DEF some

‘He says he has an animal (lit. He says himself own/be an owner of an animal)’
(Heusler 1950:126)

We will return to this interesting use of -IN and to the possible interdependence of the grammaticalization of the definite and the indefinite article in Swedish in Chapter 6.

---

64 I am thankful to Folke Josefsson for drawing my attention to the fact that some occurrences of -IN (quoted as (119)) in the OS material that I found difficult to account for might in fact be seen as marking specificity rather than definiteness.
4.6.2. Co-occurrence of -IN with strong forms of adjectives

The use of -IN as a specific marker is one of a few surprises that the speaker of MS experiences when studying its distribution in OS, particularly in period I, at a relatively early stage of the grammaticalization. Another is the possibility for marked nouns to co-occur with strong forms of adjectives.

(122) Konæ firi giær manni fællir hana luct
woman for makes man charge her closed-ST

hæræznæmpðin
jury-DEF

‘If a woman kills a man, she shall be charged by a closed jury’.

In Chapter 3.2.1 we have discussed the origins of the so-called weak adjectival inflection in Proto-Germanic and some examples of the meaningful variation between strongly and weakly inflected adjective modifying definite noun in Icelandic.

The NP luct hærædznæmpðin combines a strongly inflected adjective and a marked noun. Such a combination is no longer possible in MS; instead, the weak inflection appears obligatorily in defNPs and the strong in indefNPs (see also Appendix, notes on NP types in OS and MS). Such phrases are also infrequent in OS. Their appearance may serve as evidence that the weak adjectival inflection in period I is not just the agreement phenomenon it becomes in later stages of language development; rather, it retains some of its original meaning of ‘individualization’, bordering on definiteness.

This interpretation is further substantiated by the relatively frequent appearance of the strong form of an adjective with a possessive or a genitive (in practically all possible word orders), while in MS the weak form is obligatory in these contexts.

(123) ANP meþ konunxlíkum kærlk sinum
with royal-S love his

‘with his royal love’

(124) PAN hans siukt ben
his sick-S leg

‘his sick leg’

65 Adjective egen ‘own’ is an exception here, it is weakly inflected after a definite determiner but strongly after a possessive, compare: den egna bílen ‘the own car (=one’s own car)’ and hans egen bíl ‘his own car’.
In all four orders, there is variation between the strong and the weak form of the adjective (Delsing 1994:99-108).

It must be noted, though, that the strong inflection is much more frequent after possessives and genitives than with marked nouns. Only occasional examples of the latter have been noted.

### 4.6.3. Co-textual uses of -IN in religious prose

When discussing the unique referents and their marking in 4.5.1 we noted a number of uses of -IN which elude simple classification as non-textual. In period II, there is a striking difference between religious prose and profane narratives when it comes to functions in which -IN is found. While there are almost no instances of non-textual use in profane prose, they constitute 1/4 of all occurrences in Bo and more than half in HML! This has mainly to do with the character of the texts: Bo describes the childhood of Jesus Christ, with stories loosely based on what is known from the Bible, but giving more detail of mainly sentimental nature. References are constantly made to concepts and facts known from the Bible, and these need no textual introduction; instead they are used in the definite form from first mention. When the reader may have doubts, an explanation follows.

(127) **Nw** tha the **komo** fore **stallit** som är thz

shed which **herra** var föddir ij

**skiulit** som var lord was born in (Bo 1)

‘They now came to the stable, which is the shed that our Lord was born in’.

In the following example (128) the definite (**stalleno** ‘the stable’) could be antecedent-less if the noun was something like **hus** ‘house’, which is a natural place to live in. Here the source of definiteness must be the general knowledge that the reader is assumed to possess, concerning the place where Jesus was born.
When the fourtieth day was come according to the laws, our Lady went away from the stable together with the boy Jesus and Joseph.

-IN is frequently employed in Bo to present a new referent but one which the author considers important for the reader to be familiar with, as part of the reader’s general knowledge of the world and the Christian faith and celebrations. We saw before the case of stall ‘stable’; below a similar presentation with explanation can be found.

‘So when she came in the temple which we call Kirkiogang’.

Even though some referents are provided with an explanation, many are assumed to be familiar to the reader and no explanation is offered, with the definite form as the only mark of familiarity. These are all connected with the biblical story or with the celebrations and sacrifices made.

‘(Mary and Joseph go into a temple and buy two doves for their sacrifice). And the gospel says nothing of a lamb, since that was rich men’s sacrifice’.
If all uniques were divided into those connected with religion and others (primarily natural phenomena, like sun and moon, and names of authorities) the proportion of ‘universal’ uniques is more or less the same in Bo and HML as in other texts of period II. It is the ‘religious’ uniques that are responsible for the high percentage of uniques in the religious prose. It is interesting to note that they are not wholly disconnected from the texts, although they are not introduced (linguistically) by any other prior referents; they may be considered ‘introduced’ by the text genre—a collection of moralizing stories to be read in a convent or at a religious gathering or by a pious Christian on his or her own, but always in clear connection to religion. The definiteness of concepts such as dygden ‘the virtue’ or sanningen ‘the truth’ arises not out of the textual situation but a co-textual one—the context in which the text is produced and read. There are frequent references made to different virtues, like ödmiukt ‘humility’ and vices, like högfärdhet ‘vanity’, which are likewise regularly marked with -IN and cannot be explained on the basis of linguistic information in the texts. Neither is it easy to think of them as uniques or generics. Consider the following examples:

(131) Thu giuir var herra hugnadhin til at lätta
      you give our lord mind-DEF to to relieve

      os mz hopino at vi skulum ey
      us with hope-DEF that we shall not

      vanskas oc dröuilsin til ödmiuktnna
      run short and grief-DEF to humility-DEF-GEN

gömilse preservation (Bo 11)

‘You give your mind to the lord, to relieve us with hope so that we shall not falter and grief for the preservation of humility’.

(132) varfrw som fatikdomin älskadhe
      our-lady who poverty-DEF loved (Bo 3)

‘Our Lady, who loved poverty’.

(133) ok thikkir mik thän vara komin til höxsta oc
      and thinks me then be come to highest and

      vansammasto fulcomlikhetzsins trappo
      most beautiful perfection-GEN-DEF-GEN ladder (Bo 31)

‘And I think I am come to the highest and most beautiful ladder of perfection’.
Anomalous uses of -IN

The virtues and vices presented in (131-134) must be understood in connection with the text itself, its character and its purpose. Since Bo, as well as other religious texts, was primarily intended as moralizing reading for monks and nuns, the qualities discussed can be understood in their religious context: the hope therefore is ‘the hope of salvation’, the humility is ‘the proper/correct, Christian humility’ and the poverty is ‘the self-chosen, Christian negation of material goods’. Interestingly enough, the Christian interpretation is to be used also on negative qualities, like greed and indolence.

All these instances of -IN in religious texts can be explained neither textually nor non-textually but rather co-textually—as dependent on the purpose of the text. At the same it must be noted that MS has a tendency to mark many abstract nouns expressing emotions and qualities, where English uses a bare (unmarked) noun.

Some of these examples are proverbs or sayings, where irregular use of definite and indefinite article may be expected, but as can be seen in the (138) exam-
ple, such ‘overuse’ (from the English point of view) of the article is not limited to lexicalized phrases only.

Religious prose introduces a number of other referents absent from the legal texts and narratives. The referents are also connected with the religious character of the text and include terms such as *siāl* ‘soul’ and *licaman* ‘body’. These are typically marked with -IN without being previously introduced in the text, either directly or indirectly, which is also understandable from the MS perspective—if they are used generically.

(139) *ok alt thz som huxsins roo oc and all that which mind-GEN-DEF-GEN serenity and siāfleek hindrar fly swa som etir oc calm prevents flee so as poison and siālinna owini soul-GEN-DEF enemy* (Bo 46)

‘And everything that disturbs the peace of mind, flee from it like from poison and the soul’s enemy (=devil)’.

(140) *at han vil andans oc äkke licammans that he wants spirit-GEN-DEF and not body-GEN-DEF ökno desert* (Bo 47)

‘that he wishes for the isolation of spirit and not body’.

As we can see from the examples, there are a number of concepts relating to the inner being—apart from the Nordic *hugher*66 (or *hughnad*) there are the loans *ande* ‘spirit’ and *siāl* ‘soul’67, used more or less interchangeably in reference to the immortal part of human being, ‘soul’68, and the former used mainly in reference to ‘mind’, in which use *akt* ‘mind’ also occurs. The definite form may be

66 *hugher* (= mind) origin unknown, also used as ‘memory’
68 Both *ande* and *siāl* are used in many runic inscriptions, apparently for emphasis rather than to express different concepts, as in the inscription (U 323) quoted below:

*iystin auk iuruntr auk biurn þiR byryþr riþþu stin trums faþur sin kuþ ihlbi ons ont auk selu fur kifi onum sakaR auk sutîR hi (..)*

Östen and Jorund and Björn, these brothers raised (this stone after) their father. May God help spirit his and soul, forgive him sins.
interpreted generically or as an extended instance of inalienable possession (analogical to body parts’ terms).

Occasionally, we find -IN used generically in a way that is foreign to MS.

(141) vidhir bärg hit findir thu dalin 
by mountain-DEF find you dale-DEF (Bo 10)

‘You find a dale by every mountain (= good and bad things go hand in hand)’.

Example (141) is one where the MS would probably not use the definite, but rather Vid varje berg finner du en dal ‘By every mountain you find a dale’, with ‘every’ and indefinite rather than definite articles. Also in (142), MS would prefer the indefinite article.

(142) Än här ödhmiukadhe thu thik ok miswyrdhe 
yet here humbled you yourself and dishonoured
thik, takande vppa thik thrälsins ham ok
yourself taking upon you thrall-DEF body and
äkke konungxsins oc vardh swa som en
not king-DEF and became so as one
af os pilgrimbir oc wtläninge
of us pilgrims and foreigners (Bo 37)

‘And yet you humbled yourself and denied yourself glory, taking upon you the body of a slave and not of a king and you became one of us, pilgrims and foreigners’.

In period III, there are only few instances of what, from the MS perspective, may be termed an overuse of -IN. They include the co-textual definites that we have discussed so far. As we have noted, the MS feel of overuse is a result of the modern reader’s unfamiliarity of the texts and the conceptual world they represent, rather than a truly foreign use of the form (unlike the use of bondin ‘the yeoman’, which is indeed foreign to MS). The distribution of -IN in period III seems to be largely identical to MS.

4.6.4. Co-occurrence with numeral EN

The marked form may also co-occur with numeral en ‘one’. This is studied in greater detail in Chapter 5.9.
4.7. Discussion of the distribution of -IN

In 4.4 and 4.5 we have seen examples of -IN in different textual and non-textual functions. We have also seen that although the textual functions are already well-represented in period I, in particular in the religious prose, we do find numerous counterexamples where bare nouns are used—the usage is not yet obligatory. We have also seen that the overall frequency of -IN rises from periods I to III though it does not reach MS frequencies (Table 9). We are now going to look at the distribution of -IN in the textual and non-textual functions in each period. Table 28 shows the results for first 100 -IN forms in each text from periods I-III, apart from AVL, where only 24 are found. The comparison material is the first 100 definite forms (marked with -IN) from an MS short story (Mari Jungstedt En helg på landet⁶⁹, 2006).

Table 28. Distribution of -IN in textual and non-textual functions in OS

<table>
<thead>
<tr>
<th>Period</th>
<th>Textual functions</th>
<th>Non-textual functions</th>
<th>Other*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Direct anaphora</td>
<td>Indirect anaphora</td>
<td>Unique reference</td>
</tr>
<tr>
<td>Period I</td>
<td>66,2%</td>
<td>28,4%</td>
<td>1,9%</td>
</tr>
<tr>
<td>Period II</td>
<td>55,5%</td>
<td>12%</td>
<td>21,7%</td>
</tr>
<tr>
<td>Period III</td>
<td>48,8%</td>
<td>21,6%</td>
<td>14,2%</td>
</tr>
<tr>
<td>MS</td>
<td>18%</td>
<td>61%</td>
<td>17%</td>
</tr>
</tbody>
</table>

* Lexicalized phrases or NPs with nästa ‘next’, första ‘first’ etc.

In period I -IN is overwhelmingly used in textual functions with direct anaphora dominating heavily. This is particularly striking as only 2 out of 6 texts studied include stories where anaphoric reference is to be expected—4 texts are laws. In spite of the lack of narratives, the percentage of direct anaphora is over 66%, suggesting that this is the original function of -IN.

Proportions change gradually, although the textual use still dominates over the non-textual one, where generics come to be marked with -IN in period III.

Definite plurals lag behind definite singualrs, as can also be seen in table 29. The proportions are close to MS in period III.

Table 29. Singular and plural -IN phrases in OS

<table>
<thead>
<tr>
<th>Period</th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period I</td>
<td>95,4</td>
<td>4,6</td>
</tr>
<tr>
<td>Period II</td>
<td>90,3</td>
<td>9,7</td>
</tr>
<tr>
<td>Period III</td>
<td>85,8</td>
<td>14,2</td>
</tr>
<tr>
<td>MS</td>
<td>84</td>
<td>16</td>
</tr>
</tbody>
</table>

⁶⁹ Three MS short stories were studied to compare the frequencies of -IN and EN as well as their distribution in MS with that in OS. Of these three, the one by Mari Jungstedt had the most ‘average’ results, therefore it was chosen as comparative material for this part of the study.
Grammatical functions of defNPs in OS show surprising stability. Subject dominates slightly in periods I and II and defPPs rise in number in periods II and III, but the figures show neither radical changes nor clear tendencies.

Table 30. Grammatical functions of -IN phrases in OS

<table>
<thead>
<tr>
<th>Period</th>
<th>Subject</th>
<th>Object</th>
<th>PP</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period I</td>
<td>32,8</td>
<td>26,6</td>
<td>30</td>
<td>10,6</td>
</tr>
<tr>
<td>Period II</td>
<td>36,2</td>
<td>15,7</td>
<td>38</td>
<td>10,1</td>
</tr>
<tr>
<td>Period III</td>
<td>26,8</td>
<td>20,6</td>
<td>33,6</td>
<td>19</td>
</tr>
<tr>
<td>MS</td>
<td>31</td>
<td>14</td>
<td>50</td>
<td>5</td>
</tr>
</tbody>
</table>

In comparison with MS figures we may note that the percentage of objects has fallen while the PPs have gained ground. However, this is mainly dependent on a different change from OS to MS, namely the loss of the case system: where OS shows direct or indirect objects, MS instead has prepositional objects in many cases.

As can be seen in Table 30, a number of uses of -IN can be classified as adverbials (without preposition, e.g. *hela natten* ‘the whole night’) or attributes.

(143)  
A thæssum daghum a bonde oc husfru ofra (.)  
on these days have peasant and wife offer  
varafru dagh þæn första vm hösten  
ourlady day this first on autumn-DEF (YVL KB:54)

‘On these days shall sacrifice be made: (.) Our Lady’s day, the first of autumn days’.

Frequency is in MS expressed by means of a definite: *om dagen* ‘a day’, *om året* ‘a year’, *i månaden* ‘a month’. Further, some -IN occurrences are due to the appearance of certains forms, such as *samma* ‘same’, *nästa* ‘next’, *första* ‘first’.

(144)  
þa skal han i dögrenoste första oc  
than shall he in day-and-night-DAT-DEF first and  
daghinom næsta soknenne buth fa  
day-DAT-DEF next parish-DAT-DEF message get  
(YVL KB:3)

‘Than he shall inform the parish within two days and night’.

---

70 There is no single lexeme in English corresponding to the Swedish *dygn* ‘day and night, 24 hours’.
4.8. Conclusions

It has been shown here that although -IN appears in both textual and non-textual uses already in the oldest text, AVL, it is nevertheless possible to study its spread from strictly textual functions (direct anaphora) through indirect anaphora to unique reference and—only as the final step—to generic contexts, in which it appears only sporadically as late as period III. The textual functions, mainly direct anaphora, dominate heavily in period I. Proportions between textual and non-textual functions of -IN similar to those in MS are first found in period III. Interestingly enough, the majority of textual uses are constituted by indirect anaphora, which rely on knowledge outside the text. The original function of textual deixis is reduced (as other elements may appear in this context, mainly personal and demonstrative pronouns). Factors involved in the grammaticalization include text type, not only in the traditional sense of stylistics but also in the sense of the purpose of the text: some definite uniques are closely connected with the religious prose showing a co-textual character.

Table 31 presents the stages of grammaticalization of the definite article in Swedish based on the OS material. A difference was made between ‘spread to context’, which is to be understood as first, possibly sporadic, occurrences of -IN in a given function and ‘obligatorification in context’.

<table>
<thead>
<tr>
<th>Stages of grammaticalization of the definite article</th>
<th>Functions</th>
<th>Spread to context</th>
<th>Obligatorification in context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Period I</td>
<td>Period II</td>
</tr>
<tr>
<td>Stage II</td>
<td>direct anaphoric marker</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Stage III</td>
<td>indirect anaphoric marker</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Stage IV</td>
<td>uniqueness marker</td>
<td>(+)</td>
<td>–</td>
</tr>
<tr>
<td>Stage V</td>
<td>generalized article</td>
<td>–</td>
<td>(+)</td>
</tr>
</tbody>
</table>

We have seen that -IN is found in practically all contexts as early as in period II, but its obligatorification as uniqueness marker is completed first in period III. Even then, it is not yet regularly used in generic contexts.

In Chapter 5 we will study the development of the indefinite article, which shows some similarities to that of the definite. In Chapter 6 we will try to establish a connection between the two grammaticalizations.
5 Grammaticalization of the indefinite article

5.1. Introduction

As noted earlier in this study, the indefinite article in MS has grammaticalized from a unitary cardinal, which is cross-linguistically the most common source of the indefinite. While in some article languages the grammaticalization has led to a new form distinct from its source (e.g. English *one* vs. *a*), the Swedish indefinite article *EN* is (still) formally identical to the numeral ‘one’. Therefore, the distinction between a numeral and an article may be challenged, just as in other languages with similar formal identity of the two, such as Spanish (see Pozas Loyo 2010 for an overview).

Since the two items—numeral and article—are identical, there is in the texts no clear dividing line between them in terms of spelling (as again, there is with English *one*—*a*). In consequence, the present study is based solely on the analysis of the functions of EN. These functions include those of the original numeral, presentative marker, and specific and nonspecific markers (see Chapter 2). They are presented in chronological order of development. It must be noted though that it is not always possible to establish with certainty the function of EN. Depending on the context, it may have different interpretations in certain instances. Drawing a sharp line between different functions is not always possible due to the limited context.

---

71 One difference may be that the article is never stressed, while a numeral may be, e.g. in contrastive use such as:

Det var EN student som inte kom, inte tre.

‘It was ONE student that didn’t show up, not three’.

However, numerals need not be stressed either, and are not if no contrast is intended. Therefore, formally, the numeral and the article are identical.

72 E.g. Terner 1922 treats EN in MS as a numeral with ‘secondary uses’ in indefinite article functions.
5.2. General results

In OS, EN could be used both adnominally and prenominally (as it is in MS). Since it is the adnominal occurrences that form a context for the grammaticalization of the indefinite article, only these were considered in the frequency count.

As we have seen in chapters 2 and 4, grammaticalization of an item is often accompanied by a rise in frequency of its use, especially at the inception of the development. An analysis of the frequency of EN in consecutive periods confirms this correlation.

Table 32. Occurrences of EN (adnominal) in Period III

<table>
<thead>
<tr>
<th>EN</th>
<th>Period I</th>
<th>Period II</th>
<th>Period III</th>
<th>MS narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>occurrences per 1000 words</td>
<td>3,82</td>
<td>8,7</td>
<td>9,56</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 32 shows that the frequency of EN rises significantly between periods I and II when its values are more than doubled and slightly between periods II and III. In period III it still does not reach the MS values.

A detailed look at the texts (table 33) reveals differences between them, which are partly genre-dependent. Legal prose shows, quite like the runic material, low frequencies and only cardinal uses of EN—which is hardly surprising, as we would expect more numerals in a legal text then in a narrative (cf. Pozas Loyo 2010:148). For each period, the texts with the highest and the lowest values respectively are marked in bold.

Table 33. Occurrences of EN throughout the texts in Periods I-III

<table>
<thead>
<tr>
<th>Period</th>
<th>Text</th>
<th>Length of the text (no of words)</th>
<th>Total no of occurrences (adnominal)</th>
<th>Per 1000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>AVL$^{73}$</td>
<td>15000</td>
<td>35</td>
<td>2,34</td>
</tr>
<tr>
<td></td>
<td>YVL</td>
<td>9200</td>
<td>27</td>
<td>2,93</td>
</tr>
<tr>
<td></td>
<td>OgL</td>
<td>10220</td>
<td>23</td>
<td>2,25</td>
</tr>
<tr>
<td></td>
<td>DL</td>
<td>10750</td>
<td>62</td>
<td>5,77</td>
</tr>
<tr>
<td></td>
<td>Bur</td>
<td>9770</td>
<td>60</td>
<td>6,14</td>
</tr>
<tr>
<td></td>
<td>Pent</td>
<td>10077</td>
<td>35</td>
<td>3,47</td>
</tr>
</tbody>
</table>

---

$^{73}$ Excerpt of AVL was further complicated by the homographic EN meaning ‘if’, which may appear in contexts where indefinite article is conceivable, e.g.
In period I there is little difference between the genres and the legal and religious prose are not very different from each other, as is the case with the frequency of -IN. But stands somewhat apart with DL also showing a relatively high frequency of EN. The low frequencies of EN in all period I texts also coincide with a low level of grammaticalization, as we will see in the rest of the chapter.

Interestingly enough discrepancies between the texts are largest in period II. In periods I and III EN shows similar frequencies across texts, but in period II Jart has 9 times as many EN as Bo! In period II Jart clearly stands out, with values double that of the nearest, SVM. Additionally, there is great discrepancy between HML and Jart, though both represent the same genre and are found in one manuscript today (dated at ca 1385, though it is likely that HML is older than Jart, see Chapter 1). The nature of both texts is similar, though the stories in Jart are decidedly shorter, more adventurous and display more characters than those in HML.

<table>
<thead>
<tr>
<th>Period II</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>HML⁷⁴</td>
<td>10000</td>
<td>53</td>
<td>5,3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jart</td>
<td>10623</td>
<td>240</td>
<td>22,6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>16350</td>
<td>169</td>
<td>10,34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KM</td>
<td>10940</td>
<td>68</td>
<td>6,2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bo</td>
<td>14300</td>
<td>36</td>
<td>2,51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JB</td>
<td>2400</td>
<td>13</td>
<td>5,41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Period III</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>PK</td>
<td>4728</td>
<td>46</td>
<td>9,73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did</td>
<td>10945</td>
<td>181</td>
<td>16,54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linc NR</td>
<td>9835</td>
<td>85</td>
<td>8,64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linc JB</td>
<td>11088</td>
<td>58</td>
<td>5,23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Troj</td>
<td>9780</td>
<td>75</td>
<td>7,67</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In period I, the table shows the following:

<table>
<thead>
<tr>
<th>EN</th>
<th>maðær</th>
<th>krævær</th>
<th>man</th>
<th>skuld</th>
</tr>
</thead>
</table>

If man demand man debt

(ANDL RB:7)

‘If a man demands (repayment of) debt from another’.

⁷⁴ Reader familiar with Brandtler and Delsing 2010 will notice that the result per 1000 words here is only half of what they give (p. 30). They find 278 occurrences of en in the entire text, but my search does not confirm it—there are 258, but of everything that one gets when searching for ‘_en’, ‘_een’, ‘_eet’ (‘_et’ returns only Latin examples) and in these results a number of other frequent words beginning with en are included, e.g. ensampnin ‘alone’. Further, they do not seem to take notice of whether en appears pronominally or adnominally. Here only adnominal occurrences are considered and the concordance has been manually pruned of all words other than the inflected forms of EN.
All these factors account for the striking discrepancy between the frequencies of both articles.

In period III the text with the highest frequency of EN is the narrative Did, closely followed by PK. The text with lowest values is Linc JB, which is different from the others in many respects (see Chapter 1).

In the rest of the chapter we will look at the original cardinal uses of EN in period I as well as at its rise as a presentative marker. In connection with this we will discuss the notion of topic and the role of initial subject in establishing topics. In 5.5 we turn to the specific uses of EN and the difference between them and the presentative uses. In 5.6 we discuss the non-specific occurrences of EN. This is followed by a discussion of the variation between EN and NGN in the non-specific contexts. Other competing markers, such as -IN and ANNAR are also discussed. The chapter ends with a summary and concluding remarks.

5.3. Numeral EN

In the runic material adnominal EN is used exclusively as a numeral (see Chapter 3.4). By common agreement no instance of EN in the oldest OS text, AVL, can be regarded as an article (e.g. Brandtler and Delsing 2010:29). The form has also a low frequency (see 5.2). Its numeral meaning can usually be deduced from a contrasting numeral used in the context (examples (1)-(3)).

(1) Drukknæ tuer baþir a enu skipi hvarghi
  drown two both on EN ship each

þærra ær annars arve
they-GEN is other’s heir

(AVL ÆB:13)

‘If two men drown on one ship, they are each the other’s heir’.

(2) Sva ær firi gipt at skiliæ þaghær þer
  so is for marriage to understand from they

kumæ baþi a en bulstær ok vnþir ena bleo
  came both on EN bed and under EN sheet

þa a hvn þridiungh i bo
then has she third-part in bo

(AVL GB:9)

‘So is marriage to be marked: from the day they both came on one bed and under one sheet then she is entitled to a third of the estate’.
In examples (1)-(3), and in the majority of other EN-instances in AVL, there is an explicit contrast between EN and some other number (two brothers—one woman, two men—one ship). It is clear that EN in AVL is a numeral and used to signal singularity. Even in the absence of contrast the numeral function is still clear, as in (4), since the fine to be paid would certainly be higher if more than one tree was destroyed.

Leijström 1934, in the only relatively more comprehensive study of the formation of the indefinite article in OS\textsuperscript{75}, notes a few examples from the legal prose, where he proposes specific rather than numeral reading of *en*, e.g.

\textsuperscript{75} The study is a chapter in his thesis on indefinite article (or rather article-like uses of the numeral) in Icelandic.
‘Now a man can charge another on one ting (extraordinary, gathered for this particular purpose) and summon him again, with two fifths present. If he returns and summons him again he may not do this until the first charge is examined for which he summoned the first time. (If the accused is found guilty) he shall pay three marks for each ting’.

It seems that Leijström’s interpretation is too radical here, since there is a clear contrast between one man, who is to be sued and two different lawsuits against him.

5.4. Some remarks on specific and nonspecific uses of EN

In Chapter 2.5.2 we have presented Heine’s model of grammaticalization of the indefinite article, repeated here as (6).

(6)  Stage I: the numeral ‘one’
Stage II: the presentative marker
Stage III: the specific marker
Stage IV: the nonspecific marker
Stage V: the generalized article

(Heine 1997:72-3)

In an empirical study, there arises a practical problem of differentiating between the functions specified in the model, in particular between the specific and the nonspecific marker. The common ‘definition’ of specificity is based on the speaker’s intentions: interpretation is specific if the speaker has a referent in mind (= a certain), nonspecific if he does not (= any). Speaker intentions are nigh on impossible to study in a historical text, therefore a more practical tool is necessary.

There are a number of contexts—so-called negative polarity contexts, that induce opacity in the interpretation of the indefNP. Consider (7)76.

(7) Mary wanted to marry a millionaire.
   .. but she could not find any.
   .. but he rejected her.

Usually, if no further context is provided, the default interpretation is nonspecific.

---

76 See also Chapter 2, examples (14) and (15).
The negative polarity contexts will be studied in 5.7. There is, however, yet another strategy to avoid analysing speaker intentions (specific or nonspecific). It is based on Heim 1988. In her study of definite and indefinite NPs she proposes the notion of *lifespan* of a discourse referent. Consider first examples (8) and (9).

(8a) I have a car. It is a Saab.
(8b) I don’t have a car. *It is a Saab.
(9a) I wish I had a car. *It is a Saab.
(9b) If I had a car, it would be a Saab.
(9c) I wish I had a car. ?It would be a Saab.

In (8a) the indefNP establishes a new discourse referent, which may be taken up in later discourse. In (8b), however, the same indefNP which is under the scope of negation, establishes a discourse referent which cannot be taken up in later discourse. Similarly in (9a), though in (9b) a further reference is possible—the second mention of the referent is still under the scope of the operator. In (9c) the second mention lies outside the scope of the operator. However the example is not entirely invalid because a similar operator is used in the second sentence.77

We may thus conclude that the most important difference between the specific and nonspecific EN is that the former establishes a discourse referent of (at least potentially) considerable lifespan, while the latter establishes a short-lived discourse referent which cannot be referred to outside the scope of the operator under which it is introduced.

The notion of lifespan will be of further use when we try to differentiate between the presentative and specific functions of EN.

5.5. Presentative EN

The use of ‘one’ to mark new discourse referents is its first non-numerical function and marks the inception of the grammaticalization process (see Chapter 2). A characteristic feature of this use of EN is that the referent introduced by an EN-phrase is picked up in later discourse—the original presentative function is thus limited to prominent referents with long lifespan78 (see 5.4), which show discourse *persistence* (Givón 1983).

---

77 Possibly there is an ellipsis of (If I had one) in the second sentence.
78 In other words, a new discourse referent established by an EN-phrase cannot be abandoned; in this respect EN at the beginning of grammaticalization is similar to English ‘this’ used to present new discourse referents.
In MS, new discourse referents are typically presented by means of direct objects (DO) located postverbally. Consider the following example:

(10) Han har köpt en bil. **Den** är röd.
‘He has bought a car. **It** is red.’

Since the original function of the indefinite article is to present new discourse referents and since that is normally done by means of objects, we would expect to find high frequency of EN-objects in the oldest Swedish material. However, an overview of the grammatical functions of the EN-phrases (diagram 6) does not confirm this hypothesis. The dominant function of EN-phrases in periods I and II is instead subjects.

Diagram 6. Grammatical functions of EN-phrases in OS, periods I-III. [\%]

![Diagram showing the percentage of EN-subjects, EN-objects, and EN-PPs in periods I, II, and III.]

Diagram 6 shows that in periods I and II the dominant grammatical function of EN-phrases is not the object but the subject. The situation changes first in period III, when the percentage of EN-subjects among EN-phrases falls to little over 12% and is lower than the percentage of either objects or PPs\(^{79}\).

We can see in Diagram 7 that in Bur, the oldest text where EN is used in functions other than the original numeral one, in the religious texts HML and Jart, as well as in the narrative KM, it is not the objects but the subjects that dominate among EN-phrases. In Pent, objects are slightly more frequent than subjects but it is first in the younger texts, notably in period III (PK, Did, Linc NR and Linc JB), that objects start dominating over other grammatical functions of EN-phrases.

\(^{79}\) Grammatical functions not included in the diagram are: adverbials, attributes, predicatives, appositions.
It seems therefore that the presentative role is fulfilled mainly by subjects and not by objects in periods I and II. In MS EN-subjects are mainly found in non-initial positions: consider examples (11a-c).

(11a)  ?En katt satt i trappan.

‘A cat sat on the stairs’.

(11b)  I trappan satt en katt.

‘On the stairs (there) sat a cat’.

(11c)  Det satt en katt i trappan.

‘There was a cat (sitting) on the stairs.

MS avoids sentence-initial indefinite subjects by either placing adverbials initially (11b) or by introducing a dummy subject det (11c). The rare sentence-initial indefinite subjects that can be found in MS prose are not text-initial, as in (12a-b). In fact, they may be considered a type of ‘indirect anaphora’, since they are also determined by an antecedent. It is impossible to treat this interesting phenomenon exhaustively here, suffice it to say that indefNPs are occasionally found in same contexts as defNPs, with an antecedent ‘anchoring’ them in the text. The difference, however, is that although they are connected with a prior information, they are not unique in their reference. On indefNPs with antecedents, see e.g. Fraurud 1990:404-405, Wald 1983 and Ushie 1986.
(12a) Foten var värre. En lukt började kännas bara efter någon dag. ‘The foot was worse. A smell started to be felt just after one day’.

(12b) Hon tog rast och satte sig på kälken. En kort stund bara, hon var blöt av svett och skulle snart frysa. En ekorre sprang över snön precis intill, såg henne inte. ‘She took a break and sat on the sledge. A short moment only, she was wet of sweat and would soon freeze. A squirrel ran over the snow next to her, did not see her’.

(Larsson 2006:203 and 205)

In OS, on the other hand, the sentence-initial EN-subjects are also text-initial and antecedentless.

(13) En vælbúrin ungar suen forløet værudena ok
 EN well-born young boy left world-DEF and

folgþe andream: hans frændar vildo þom baba inne brænna
 followed Andreas his relatives wanted them both in burn

ok þendo brand iui þera hærbærghe Smasuennen
 and lit fire over their shelter small-boy-DEF

slækte balet mz litlo vatne þa vildo þe
 quenched fire with some water then they

klíva í huset ok wrÞo alle iæmskyt
 step in house-DEF and became all suddenly

blínde: þa ðpte en at allom: vae þer striþa
 blind then called one to all danger is fight

vþ guþ som sæluan haldar mz andrea.
 with God which self holds with Andreas

ok þom hanom fólgrer. hær af worþo mange
 and them him follow this of many

christne. ok troþo þy andreas predicaþe
 Christian and believed this Andreas preached

‘A well-born young man left the world and followed Andreas. His relatives wanted to burn them (alive) and started a fire over their shelter. The young man put out the fore with some water. Then they wanted to enter the house and became all suddenly blind. Then one called to all: it is dangerous to fight against God himself, who protects Andreas and all who follow him. Of this many became Christian and believed what Andreas preached’.
On the other hand, OS EN-subjects that are not text-initial are not sentence-initial either: compare the text-initial *en diäkne* ‘a deacon’ with a text-medial *en prästir* ‘a priest’ in the following example. While the former introduces the main character of the story, the latter introduces a background one.

(14) **EN diäkne war j enom stadh som heet montepessolanus hulkin som foor mz diäwlskap oc truldom. thän sami for annars sins kompans bön. the wilde wita tidhinde aff sino lande kalladhe til sin diäfwlen. ok diäfwlen vppinbaradhis j enna quinno liknilse ok syntis hanum idhkelika j tholiko liknilse Diäknin iäfwadhe hwas thät war hälldir diäfwl hällas ey. ok tho spordhe han at diäwlin sit ärande ok han swaradhe hanum. Nu j the stundinne for ther fram vm en prästir mz gudz likama til en siukan man ok klokkad ringde for gudz likama. Ok ginstan wiste diäfwlen at thät war sandir gudh som prästin fördhe Ok ginstan j samu stundinne fiol han nidhir a iordhina a knä ok negh gudz likama Diäknin sa thetta ok sagdhe Nu tro iak sannelika at thz är ey diäfwl som thu kalladhe, for thy at han hafðhe ey fallit a knä som han giordhe Tha swaradhe diäfwlen ok sagdhe hanum. West thu ey hwat sciwat star j the hälgho skript som sighir J ihesu christi hedhir ok namn skulu böghias al knä. badhe j himerike ok iordhrike ok hälwite ok ginstan war han borto fran them.**

‘There was a deacon in a town called Montepessolanus, who practiced witchcraft, he called the devil asked by his friends, who wanted to know news from their land. And the devil revealed himself in a woman’s image and always appeared in such image. The deacon hesitated whether it was the devil or not and asked what he (the deacon) wanted of him (the devil) and he answered him. In this moment a priest went by with God’s body (travelling) to a sick man and the bell rang for the God’s body. And immediately the devil knew it was the true God that the priest carried. And immediately in the same moment he fell down on earth on his knees and bend for the body of God. The deacon saw this and said: Now I truly believe you are not the devil as you claimed, because he would not have fallen on his knees. Then the devil answered and told him: Do you not know what is written in the holy script which says that in the name of Jesus shall each knee bend, both in heaven and on earth and in hell. And immediately he was gone from them.’

It seems that OS allows initial EN-subjects on a greater scale than MS does. In fact, in period I, the initial EN-subjects dominate over the non-initial ones.

Diagram 8 shows the proportions between initial and non-initial EN-subjects in periods I-III. As can be seen, the initial subjects dominate over the non-initial ones in period I and their frequencies are similar in period II, though here the non-initial EN-subjects start to dominate. It is first in period III that initial EN-subjects diminish.

We could describe the difference between the two types of EN-subjects as pertaining to the character of the topic introduced by each of them. While the text- and sentence-initial EN-subjects introduce a text-topic, the text-medial EN-subject introduces a sentence-topic.
Topic is different if seen from the sentence (or clause) perspective than if seen from a paragraph/text or discourse perspective. The topic of a clause may, but not necessarily, be identical with the topic of the paragraph or the topic of the discourse. An example of this differentiation is provided by some of the OS texts studied. E.g. HML (*Helgamanna leverne*) is a collection of hagiographic tales, as its title suggests: ‘The lives of holy men’. We may say that the discourse-topic is a group of people, considered to be moral paragons, and their lives. When we look to individual stories we note that they sometimes start with a title, e.g. *Af enom munke* ‘about a monk’ and start with a sentence-initial EN-subject, e.g.

(15)    En munk kom.
       EN monk came

Such sentence-initial EN-subjects do not introduce a discourse referent that is wholly new, rather specify a potential one. Another type of such ‘false’ introduction is found in SVM (*The seven sages*). In this text there is one superior story of an emperor whose son had been away to study under the tutelage of seven wise men. The emperor has a new wife who upon the son’s return incites the emperor to have him executed. She strengthens her demand by telling the emperor a moral tale. This tale is countered by one told by one of the seven sages. The wife and the sages continue to warn the emperor of what the future might hold for him with either course of action—to kill or spare his son. Many of the stories start with a sentence- and text-initial EN-subject but for each of them the ground is prepared in the earlier text. Consider the following example.
Ok fordarffwin i idhar son tha händer idher som enom byman hände här i stadhin aff sinne skate Kesarin spordhe hurw honum hände Kato swaradh Wilin i fordragha idher sons dödh tha wil iak idher thz sighia Tha sagdhe kesaren Jak wil han gerna fordragha til thäs thu haffwer thz saght Kesarins son leddis igen i mörkastoffwona. oc mästarin byriadhe sit äwintyr.

En byman war i thenna stadh hwilkin som haffdhe ena skato.

'Sand if you ruin your son same thing will happen to you as happened to a man here in the town of his magpie. The emperor asked what had happened to him. Kato answered: If you wish to delay your son’s death then I will tell you. The emperor said: I will gladly delay until you have told me the story. The emperor’s son was led again to prison and the sage began his story.

A man was in this town who had a magpie.'

The initial EN-subject establishes the main topic of the story (which is normally only one or two paragraphs long). This text-topic continues over all sentences (clauses) and is picked up by either pronouns, definite NPs or, in cases of maximum accessibility, by zero anaphora. Other, more ‘local’ topics may be established as well, though usually through non-initial subjects or objects rather than sentence-initial subjects (see example (14)).

It has been claimed in literature that the initial position in the clause is reserved for ‘old’ or ‘given’ information, and that topics, usually initial, are always definite. These claims are confirmed in studies of the grammaticalization of the indefinite article in Spanish (e.g. Pozas Loyo 2010), where the presentative function of UN is first found with direct objects, placed postverbally.

The results from OS, however, show a different pattern. Far from establishing topics by postverbal direct objects, texts from periods I and II use dominantly initial EN-subjects for this purpose. This leads one to question whether these instances of EN can be considered indefinite (as we would interpret them today, with MS system in mind) or whether they are something else, and in this case, what that is.

Let us first consider the number of further references made to a discourse-referent established by EN-subjects. New referents presented by EN-phrases in subject positions can be taken up later in the text. The more numerous these further mentions, the more persistent the referent. A comparison between initial and non-initial subjects shows that the former are picked up in later discourse more often than the latter. Initial EN-subjects introduce text-topics, which is clear from the number of subsequent mentions of the referent.

As can be seen in diagram 9, initial EN-subjects introduce referents that are more persistent (are mentioned more times in the texts) than those introduced by non-initial EN-subjects. In period I, a discourse referent established by a sentence-
Diagram 9. Average number of further mentions of a referents introduced by EN-subject in OS

initial EN-subject is taken up 6 times on average in the following discourse, while non-initial EN-subjects establish discourse referents that are taken up only twice on average in the later discourse. This tendency is particularly clear in texts with a high percentage of the initial EN-subjects, namely religious prose of periods I and II. Coincidence of EN and the initial position is seen also in period III texts, PK and Did. In MS indefinite initial subjects are rare, and a structure Det var en.. ‘There was a ..’ is preferred. Aggregating the results for each period allows us to see the more general tendencies. For each text, the numbers may vary dependent on other factors, such as the average length of the story where the sentence-initial EN-subject is found.

These results can be interpreted in the following way: to begin with, in period I and early parts of period II, EN is used in the subject position, to present a new referent (presentative function). The use of EN is combined with initial position, giving the new referent strong text-topicality. In late period II EN is still overwhelmingly found in the subject position, but the percentage of initial subjects has dropped dramatically. In period III, EN-phrases as subjects fall in frequency and are overwhelmingly non-initial. EN still introduces new referents to the discourse but with low text-topicality.

What of the direct object, which is the ‘natural way’ of introducing topics in MS? In diagram 10 we see a comparison between the average number of further mentions of a referent introduced by subjects and objects. It is clear that referents introduced by the objects are not as persistent as those introduced by subjects.
In period III the EN-phrases, not necessarily subjects, introduce referents that may be taken up in later discourse but are not topical.

(17) Birka han laa vppa ena öö j mälar som
Birka he lay upon EN island in Mälar which

hether birköö
is called Birkö (PK 226)

'He grounded Birka (the town) on an island in the Mälar lake which is called Björkö'.

(18) Then gudhelike oc mäktoghe herran, sancta
this godly and mighty lord saint

amalbärgha fadher, haffdhe widh sit palacium
Amalberga’s father had by his palace

en widhan oc wällustoghan trägardh (..) oc
EN wide and delightful garden and

war thän trägardhen swa lusteliken at
was this garden so pleasant that

ythermeer wällustoghet kwnne man ey finna
more delight could man not find

(Linc NR 279)

'The godly and mighty lord, Saint Amalberga’s father, had by his palace a wide and delightful garden (..) and the garden was so pleasant that more delight could not be found'.
Also, the text-initial EN-subjects decrease in period III, benefitting the non-initial ones.

(19) Oppa then tidh war eth closter i beyla, iomffru upon this time was EN nunnery in Beyla virgin

   maria til hedher opbygth, i thz closter war en
   Mary to honour built in this nunnery was EN

   hælogh abbatissa
   holy abbess

   ‘In this time there was a nunnery in Beyla, consecrated to the glory of the Virgin Mary. In this nunnery there was a holy abbess’.

To sum up, the presentative function of EN in period I is realized by sentence- and text-initial EN-subjects, which establish discourse referents of a long lifespan (ca. 6 further mentions). In period II the major presentative devices are still predominantly initial EN-subjects, though non-initial subjects also show a rise in frequency as well as establishing discourse referents of a long lifespan (ca 4 further mentions). In period III EN-phrases serving as subjects (both initial and non-initial), objects and prepositional objects establish discourse referents with similar lifespan—of ca 2 further mentions, much shorter than either period I or II.

5.6. Specific EN

Heim’s notion of lifespan of a discourse referent can be used to define the difference between presentative and specific markers. The main difference between the presentative and the specific marker is that the former introduces a new discourse referent which is taken up later (persistence) while the latter introduces a new discourse referent which may, but need not, be referred to again. Needless to say, the distinction is not always clear. In example (20) two new referents are introduced: ‘en stad’ a town and ‘en kättare’ a heretic.

(20) Jac fan i enom stadh een kättara swikande
I found in EN town EN heretic fooling

   folkit
   people-DEF

   (HML 191)

   ‘I found a heretic fooling people in one place’.

The story continues with 9 further mentions of the heretic and 1 of the town. Both instances of EN may be considered as presentative markers, though there is a difference in the prominence of the discourse referents in the discourse.

We have previously seen that EN is used exclusively as a numeral in AVL but that it is nigh on obligatory as presentation marker in Bur wherein it has reached the first stage of grammaticalization. We will now consider its invasion in the next context—specific reference, which we identified in Chapter 2 as a bridging/switch context.

Presentative uses of EN discussed in the previous section are specific in the sense that the speaker has a particular referent in mind when using EN—the presupposition of existence holds. These are different from other specific uses in that though in the latter case EN signals that the referent is new, it does not necessarily signal that it is the topic. That means that the referents of specific EN-phrases need not be picked up in later discourse and may constitute back-grounded information—new, but back-grounded.

Taking the number of further mentions into account, the specific uses of EN can be located on a continuum from EN as a presentative marker, where the new referent is also the topic with numerous later mentions at one end of the continuum, and EN-phrases which rate 0 mentions in further discourse at the other end. The presence of EN usually coincides with the grammatical functions of subject or object for referents with the most number of further mentions and of preposition object or adverbial for those with the lowest number of further mentions.

(21) Oc foor ysaac bort til landz ænda oc and went Isaac away to land-GEN end and
bygde ther j en dal lived there in EN valley (Pent 161)

‘And Isaac went away to the end of the country and lived there in a valley’.

EN-phrase in this example is a PP (and PPs rank very low in further mentions). The referent en dal ‘a valley’ is not mentioned again; clearly it does not constitute the topic of either the text or the sentence but only gives the background information.

We saw in the previous section that the initial position of the EN-subject coincides in periods I and II with a high number of further mentions of the referent (and there is no instance without a further mention). When it comes to non-initial subjects, the numbers are still high, but this is not so with objects.

(22) Nw mædhan han soffuir tha syntis honom en stighi now while he sleeps than appears him EN ladder
In his sleep he sees a ladder extending from the earth and to the heavens.

It cannot be claimed that EN has gained ground as a specific marker even in the linguistically more modern texts from period I, namely Bur and Pent. Although it can occasionally be found in such contexts even alongside abstract concepts (see examples (23) and (24)), more often than not it is only the bare noun that can be found, even with concrete terms (example (25)).

(23) 
```
jak  hafuer  giort  ena  wenlika  sæmio  mædh  minom
I have  made  EN friendly agreement  with  my
modherbrodher
uncle
```

‘My uncle and I have reached a friendly agreement’.

(24) 
```
oc  var  thz  oc  en  hæmd  iffwr  fadherin
and  was  this  and  EN  revenge  over  father-DEF
```

‘And it was also a revenge upon the father’.

(25) 
```
Nw  kombir  ioseph  fram  til  sinna  brødhra  oc  æmskøt
now  comes  Joseph  forward  to  his  brothers  and  at once
føra  the  han  aff  hans  kledhom  oc
take  them  of  his  clothes  and
kastado  han  liffuandis  j  vatnlosan  brwn  æpthir
threw  him  alive  in  waterless  well  after
rubens  radhe
Ruben-GEN advice
```

‘Now Joseph comes forward to his brothers and they at once take his clothes off his back and threw him into a waterless well, still alive, as Ruben advised’.

The following quote is a good example of a bridging context, where specificity is the intention rather than information on the number of chairs the king was sitting on; as one person usually occupies one chair at a time, this information is redundant. As it is a referent not taken up in later discourse, it is not presentation either.
Similarly in the following two examples, where the intention is specificity rather than numerical or presentative information (though the box is referred to again in a later part of the story, but as an object):

(27) Tha lot k. m. k. göra et skriin aff gul ok
   then let KM make EN box of gold and
   binda mz silff bandom
   bind with silver bands-DAT

‘Then Charlemagne let a box of gold be made and bound with silver ribbons’.

(28) Sidhan war han i patriarchans borgh fyra manadha
    later was he in patriarch-GEN castle four months
    oc lot göra ena kirkio the som heter
    and let make EN church this which is-called
    sancte marie latine
    st. Mary Latin

‘He speant four months in the patriarch’s castle and let build a church, the one called St Mary Latin’.

We saw in Chapter 4 that in the early stages of grammaticalization of the definite article a number of restrictions applied to its distribution limiting its use by, among others, plural or mass nouns. EN is not found with plural referents at all, its singularity making it impossible (unlike e.g. its Spanish counterpart UN—UNOS), but its use by mass nouns is similarly limited in the early stages of grammaticalization.

(29) huilken förra war strangher som hardasta iärn
    which earlier was unyielding as hardest iron

‘The former of which was as unyielding as the hardest iron’.

This restriction applies even to mass nouns used individually (fire—a fire):
With countable specifics, EN is regularly found already in period II.

(31) Een munkir ny kumin i ödhkninna hafðhe ey
    one monk new come in desert-DEF had not

sälla Annar munkir hafðhe een toman sälla.
    cell other monk had EN empty cell

‘A monk, newly arrived in the desert, had no cell. Another monk had an empty cell’.

Summing up, EN is only occasionally used in what we identify here as specific contexts in period I, though it is regular as a presentative marker for new topics. The situation changes in period II when EN is regularly used as a specific marker by all new discourse referents, including those presented in PPs or adverbials. It is however not obligatory in the so-called negative polarity contexts— as example (31) shows, the specific ‘cell’ is marked, whereas the non-specific one is unmarked. In 5.7 we will look at negative polarity contexts in more detail.

5.7. Non-specific EN

The last stages of grammaticalization of EN include its use as a marker of non-specificity or, in Heim’s terminology, to establish new discourse referents with a lifespan confined to the scope of the operator under which they appear. We have in chapter 2 listed a number of such operators, or contexts, so-called negative polarity contexts, which induce opacity. And thus, objects for e.g. in scope of verbs of volition (e.g. ‘plan’, ‘intend’) may be read as either specific or non-specific:

(32) She intended to visit a nunnery..
    a. but there was none in the area.
    b. but it was too far to drive so he gave up.
We see that the interpretation of ‘a nunnery’ in (32) is dependent upon context. In historical material, we are not always so fortunate as to find such elaborations that clarify the interpretation and examples like (33) below are rare.

(33) han aktadhe at visitera eth iomffrw closter, huilket
he planned to visit EN virgin monastery which

han siålffwer fordhom haffdhe bygh oc stiktat
he himself earlier had built and grounded

(Linc NR 281)

‘He intended to visit a nunnery which he himself had founded and built’.

Since a reference is made to ‘a nunnery’ that was founded by the person planning to visit it, it is clearly specific.

For the purpose of the present study, the potential non-specific contexts have been limited to: comparisons, negation, conditionals and verbs of volition.

As we have seen in the previous section, period II is the time when EN moves up on the grammaticalization scale, securing its position as a presentation marker and by specific referents. As we will show here, it also occurs, though sporadically, in non-specific contexts. It is not until period III that EN starts appearing more regularly in nonspecific contexts.

In MS the unmarked form in at least some nonspecific contexts is NGN ‘some/any’ (see Chapter 1.5.3). Therefore in studying the non-specific contexts, we will look at both EN and NGN and their evolution. In period I it is the only marker not found in nonspecific contexts.

(34) haffwom alle en vilia oc dræpom han oc kastom
have all EN will and kill him and throw

han dødhan nidhir j nakan gamblan brwn
him dead down in some old well

(Pent 177)

‘We have all one will so we kill him and throw him dead into some old well’.

5.7.1. Comparison

In period I bare nouns are regular in comparison contexts (example 35). But there is some variation beginning here and EN occasionally does appear in comparisons (example 36), unsurprisingly, in Bur, which shows the most advanced level of grammaticalization of EN also in other respects.
(35) Tha esau hører thæsse tidhende tha wart han swa
when Esau hears these news then became he so
illa widh at han òpir oc grathir oc rotar
badly by that he screams and cries and roars
som tyur like bull

‘When Esau heard this he became so angry that he screamed and cried and roared like a bull’.

(36) huarn tima han mintes at han nekaþe sin
each time he remembered that he denied his
hærra runno hans òghon som en kalda
lord ran his eyes as EN spring

‘Each time he remembered his denial of God, his eyes ran like a water spring’.

In period II the variation between EN and bare nouns in comparison contexts shows growing preference for EN (see table 34).

<table>
<thead>
<tr>
<th>Text</th>
<th>Comparison marking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ø</td>
</tr>
<tr>
<td>HML</td>
<td>2</td>
</tr>
<tr>
<td>SVM</td>
<td>1</td>
</tr>
<tr>
<td>KM</td>
<td>5</td>
</tr>
<tr>
<td>Bo</td>
<td>1</td>
</tr>
</tbody>
</table>

The chronology of the texts coincides with a steady spread of EN in the comparison context. No EN is found in HML, though it must be noted that the context itself is infrequent. Both narrative texts show a more or less equal distribution of unmarked and marked nouns and Bo, a clear advance of marking, which may either be EN or the pronoun annar ‘other’. The examples are discussed below.

The two examples from HML are introduced by swa som ‘such as’.

(37) oc warþ swa som hæstir oc mule
and became such as horse and mule

(HML 185)
SVM, with only two instances of comparison contexts, shows EN with a countable noun and bare form of a mass noun.

(38) thera tungor waro rödha som brinnande kol
their tongues were red as burning coal

‘Their tongues were as red as burning coal’.

(39) Jak giordhe som en daare
I did as EN madman

‘I acted as a madman’.

KM may well be seen as being in the middle of the development from unmarked contexts to those marked with EN. In some examples EN is missing, whereas it can be found in others.

(40) oc iomfrun war swa skäär som roos oc lilia
and girl-DEF was so pink as rose and lily

‘And the girl was as rosy as a rose and a lily’.

(41) Nw far fram roland som et leon
now rushes forward Roland as EN lion

‘Now Roland attacks like a lion’.

(42) han skeen som en sool
he shone as EN sun

‘He shone like a sun’.

There is a relatively large variation in this context in Bo. Only one instance of a bare noun was found (given in (43) below), and 3 out of 4 EN instances also include annar ‘other’ (example (44)), which may in this context be used on its own (45)\(^8\).

(43) Skudha oc ämuäl piltin ihesum som ey än talar
see also likewise boy-DEF Jesus who not yet talks

vtan staar oc teer sik stadhughan swa som
but stands and appears self strong such as

\(^8\) The use of annar alone is impossible in MS, it can however appear alongside NGN.
‘See also how the boy Jesus, who does not yet speak, stands and has an appearance of a grown man’.

(44) Ok skodha granlica huru var herra ihesus (...) hwilar
and see carefully how our lord Jesus rests

sik swa langan tima hwaria nat i enne af
self such long time each night in one of

them threm fatiko sängomen swa ödhiuklica oc
those three poor beds-DAT-DEF so humbly and

swa snöppelica som en annar fatikir man
so shamefully as EN other poor man

af almoghanom
of country people (Bo 34)

‘Look also carefully how our lord Jesus rests a long while each night in one of these three poor beds so humbly and so shamefully (full of shame) as any other poor man of the countryfolk’.

(45) ok litit ther äptir var han vmskorin swa
and little this later was he circumcised such

som annar syndoghir man.
as other sinful man (Bo 10)

‘And some time after he was circumcised as any other sinner’.

Annar used in the comparison context may play the same role as it does in MS, where it has an almost jocular effect.

(46) och wart sidhen swa gamal ath han lagdis
and became later so old that he laid-PASS

annat sin j waggo och degdis mz en horn
other time in cradle and suckle with a horn

spinä som annat barn
udder as other child (PK 230)

‘And later he turned so old that he was laid for the second time in cradle and suckled with a horn udder as a child’.
Further there is one inexplicable example of a comparison made to some apparently familiar referent, which does not appear anywhere else in the text. It comes from a part of text describing the conversation between Mary and an angel. She is encouraged to disconnect herself from her family and people.

(47) Sit thy ensamin som turtur dwan
sit therefore alone as turtle dove (Bo 46)

‘Sit therefore as alone as a turtle dove’.

Turtle doves are earlier mentioned in the text in connection with making sacrifice, but no reference is made which could explain why a turtle dove is singled out as the comparison of loneliness, still less the form with -IN. It cannot be regarded as nonspecific though, unless the use is considered similar to the use of bondin in period I (see Chapter 4.6.1).

5.7.2. Negation

NPs in the scope of negation are always bare in period I. In period II the most regular form after negation is also a bare noun, as in (48).

(48) Han nekaþe oc sagþe sik ey hawa seet quinno
he denied and said self not have seen woman
siþan han sik i klostre innelykte
since he self in monastery in-locked (HML 184)

‘He denied and claimed that he had not seen a woman since he closed himself in the monastery’.

The only variation we find in this context in period II is between bare nouns and occasional uses of NGN. Not a single instance of EN after the negation ey has been found.

(49) En ärliken man war i thenna stadh som ey
one honest man was in this town who not
haffdhe ärffwingia oc ey husfru ther honum
had heir and not wife which him
matte barn födha
might child bear (SVM 131)

‘There was an honest man in this town who didn’t have an heir and nor did he have a wife who might give him a child’.
Examples from period III all include contexts that are a combination of negation and conditional or verb of volition. No EN was found in these either.

(50) at han ey finghe makt ower mina kropp
  that he not get power over my body

‘that he should not gain power over my body’.

(51) Tha konung karolus saa at hans astwändan syntes
  as king Karolus saw that his desire seemed
  ey wilia haffua framgangh
  not want have success

‘As king Charles saw that his desire seemed to have no success.’

5.7.3. Conditionals

Just as negations, conditionals are also unmarked in period I.

(52) hauar iak bygkt høght torn þæt ær eigh skæl
  have I built high tower this is not reason
  at iak brytar þæt atar niþar (...) hauar iak
  that I break it again down have I
  samulund grauit gópan brun huru ma iak han
  likewise dug good well how may I him
  sialuar atar tæppa
  self again seal

‘If I have built a high tower, there is no reason why I should pull it down; likewise if I have dug a fine well how can I clog it myself?’

First in period II there comes some variation between bare noun and EN.

(53) Fae keysarin mik i morghin spiwt thz
  give-COND emperor-DEF me in morning javelin DET

  som manz byrdhe är
  which man-GEN burden is

‘If the emperor would give me tomorrow a javelin that is a burden to carry for any man’.
In example (53), the NP appears in hypothetical conditional and has non-specific reference. In MS it would be marked with EN, which is not the case in OS. Now consider the following example, from the same passage of KM:

(54) Taghe kesarin i morghin fyra kliff bly.
take-CONJ emperor-DEF in morning four parts lead

oc stöpe thz wällande i en kätil
and cast it gushing in EN cauldron

eller i et kar

or in EN tub

(KM 257)

‘If the emperor took tomorrow four measures lead and cast it gushing into a cauldron or a tub’

In this passage the context is hypothetical as well, yet the referents are marked by EN. Possibly, this has to with their function in the sentence, not objects (like (53)) but adverbials (in prepositional phrases) falling outside the scope of negation.

The beginnings of variation and appearance of EN even in non-specific contexts can be seen in the following sequence of examples. They are taken from the first tale in KM. Charlemagne and his twelve companions have spent a night as guests of an eastern emperor. In the evening they got drunk and bragged about how they would prove the emperor weaker and less powerful than themselves. They did not know that he had installed a spy in one of the pillars in the hall. When the spy reported to him, the emperor became furious and declared he would have them all beheaded. When Charlemagne was told of it, he replied that all their swagger was due to excessive drink and that he believes a spy must have been located near them to report of it to the emperor (example (55)). The word niwsnare ‘spy’ is here used non-specifically. In MS it would be obligatorily marked with the indefinite article (att vi hade en spion hos oss ‘that we had a spy with us’).

(55) oc wänter iak at wi hafdhom niwsnara
and believe I that we had-PL spy

inne när os
within close us

(KM 258)

‘And I believe that we had a spy close to us’.

A little further on in the text, where an embittered Charlemagne continues his protestations, he says:
And yet, further on in the text, Charlemagne says the following to his host:

(57)  tha  satte  thu  qwarran  en  niwsnara  nær  os  
then  set  you  behind  EN  spy  close  us  

‘Then you left behind a spy with us’.

The reference is still non-specific, however EN is used. It may be due to the fact that the spy has now been debated for a while and it is not the first mention of him (yet of course the meaning is still not specific in (57)).

5.7.4. Verbs of volition

EN is not found with the so-called verbs of volition, e.g. ‘to want’ etc. in period I (58), but in period II occasional examples can be found (59).

(58)  Nw  bødh  min  haerra  mik  at  letha  sinom  syni  
now  bade  my  lord  me  to  search  his  son  

kono  
wife  

‘My lord asked me to find a wife for his son’.

(59)  Then  godhe  mannen  lydde  thera  raadh.  oc  badh  
DEM  good  man-DEF  obeyed  their  advice  and  bid  

them  sik  ena  vnga  qwinno  op  spöria  
them  him  EN  young  woman  up  ask  

‘The good man listened to their advice and asked them to find him a young woman’.

5.7.5. Predicatives

We have so far ignored the appearance of EN in predicatives. MS makes use of the indefinite article in predicatives that are descriptive, as e.g.
But predicatives that are attributive, like names of professions, appear as bare nouns.

(61)  Han är lärare.
he is teacher
‘He is a teacher’.

In this respect, MS is different from Modern English where all predicatives are indefinite.

The use of EN in predicatives is of a relatively late date. Period I texts show only sporadic NPs here. First EN-phrases can be found in period II (62), though there is variation (63).

(62)  ok sagþe cors vara eet valdelekt tekn
and said cross be EN mighty sign

(Bur 139)
‘And he said that the cross was a mighty sign’.

(63)  han war mäktogher konunger
he was mighty king
(KM 266)
‘He was a mighty king’.

In period III EN is obligatory by descriptive predicatives (see examples (64) and (65)), but occasional exceptions may be found, as in (66) and (67).

(64)  hans fadher heth iädewarda och war en
his father was called Edward and was EN

godher ryker bonde
good rich peasant
(PK 235)
‘His father was called Edward and was a good, rich peasant’.

(65)  Naar tässen hälga iomfrwn war än en lithen
when this holy virgin was still EN small

pigha
girl
(Linc NR 279)
‘When this holy virgin was still a young girl’.
5.8. EN vs NGN

At the early stages of grammaticalization of the indefinite article, there is usually a certain amount of competition from other items possible in the non-numeral contexts. One such item is OS is NGN ‘some’. We have already seen that it can appear in non-specific contexts in OS. In MS it appears in both specific and non-specific ones (see Chapter 1). We will now take a closer look at its frequency and distribution in OS.

NGN, which has a clear nonspecific etymology (‘I don’t know which’, see Chapter 3.2.4), seems from the beginning to be used mainly nonspecifically. Therefore when EN only appears in presentation and specific contexts in period I there is no conflict between the two. We see no variation here that resembles the one observed in OE between an and sum (reported in Hopper and Martin 1987, see Chapter 2.5.1). The usage is found in all texts, including the legal prose, even though the frequencies are low, much lower than those of the relatively infrequent EN.

When EN spreads through the specific contexts and starts to be used in non-specific ones, the two forms must share domain.

Frequencies of EN and NGN in period I are given in diagram 11. Only adnominal occurrences were taken into account.

The frequencies of NGN are much lower than of EN and show very stable results compared to the increase in frequency of EN, especially in period II.

The use of NGN in OS greatly resembles the MS one; the form appears mainly in negative polarity contexts, such as questions (68), conditionals (69-71), comparisons (72) and (73) and negations (74) and (75).

---

82 Compare studies of English (Hopper and Martin 1987) and Old Tuscan (Stark 2002).
Diagram 11. Frequencies of adnominal EN and NGN in OS

(68) Hwi  talar  thu  ey  eller  gör  mik  nakan
why  talk  you  not  or  do  me  NGN
hughnadh
pleasure
(SVM 116)

‘Why don’t you talk or give me some pleasure?’

(69) Löpær  dræll  bort  allær  ambut  fra  lawarþi  sinum
runs  slave  away  or  slave-woman  from  master  his
ok  gör  nokon  skaþa  dreþær  stial  allær  rænir
and  does  NGN  harm  kills  steals  or  robs
(AVL RB:11)

‘If a slave (man or woman) runs away from his/her master and does some harm: kills, steals or robs’.

(70) Vm  nakar  pelagrim  kom  in  i  stadhin
if  NGN  pilgrim  came  in  in  town-DEF
tha  skynde  hwar  for  annan  at  ledha  han
then  hurried  each  for  other  to  lead  him

Hwi  talar  thu  ey  eller  gör  mik  nakan
hughnadh
pleasure
(SVM 116)

‘If a pilgrim came to the town, everyone hurried to welcome him to his house’.
Thereafter he bade and advised her father and mother that they should send her to some holy congregation in some nunnery, which they quickly did’.

‘He asked where earth was higher than heaven (..) his (earthly) body is higher than any heaven can be’.

‘And when he came back and his mother saw him, she was happier than any man can tell’.

‘They had not a care about clothes or food or any other earthly concerns’.

none of them uplifted their wings to fly or
Examples reminiscent of the MS use of NGN include even the imperative (76-78).

(76) Giuin hælghom benom nokot ærlekt rum
give holy bones NGN dignified place (Bur 167)

‘Give the holy bones some place worthy of them’.

(77) Herra giff mik han nakan tiidh tha wil iak
lord give me him NGN time since want I
oc wäl koma han til at tala
and well come him to to speak (SVM 115)

‘Lord, let me have some time with him since I want to encourage him to speak’.

(78) Stath opp snarlika oc fly til nakra wärn mz
stand up soon and flee to NGN care with
thinne syster
your sister (Linc NR:295)

‘Stand up quickly and flee to some protection with your sister’.

In none of the negative polarity contexts exemplified above can NGN be said to be obligatory and there is a certain amount of competition between NGN and EN in conditionals (where en is the unmarked form in MS).

Naturally NGN is also used with plural referents, since EN does not appear in the plural, see (79) and (80).

(79) Sama daghin ther greffwin war dödh wordho ther
same day-DEF there count-DEF was dead were there
nakre riddara gripne oc dömpde til hängilse fore
NGN.PL knights caught and sentenced to hanging for
roff skuld plunder sake (SVM 155)

‘Same day the count died some knights were captured and sentenced to be hanged for the plunder’. 
Even though NGN seems to have been originally a purely non-specific marker, we noted in Chapter 1 that in MS it can be used in clearly specific contexts with a near-numeral meaning (corresponding to ‘more or less one’) or when the speaker is unwilling or unable to provide the hearer with more exact information, or when the information is deemed irrelevant (vague). As noted in Nivre 2002 such uses of NGN are foreign to other Scandinavian languages and are either part of a common heritage that was lost in other languages but retained in Swedish, or are a Swedish innovation. The present study is only concerned with the development of NGN as far as the overlap with EN is concerned, but it can be noted that these two uses of MS are not found in period I, but are noted in periods II and III (see examples (81-84), which indicates that they may be a Swedish innovation. However, a systematic study into both OS and Old Danish is necessary to validate this claim. The examples noted from OS are given below.

(81) the gingo til en stadh som kallas ermepolim
they went to EN town which is-called Ermepoli
oc ther legdho the sik nakat litit hws oc
and there rented they self NGN small house and
bodho ther siw aar
dwelled there seven years (Bo 14)

‘They went to a town called Ermepoli and rented there some small house and lived there for seven years’.

(82) The platadho stafuin vidh hans graf hwilkin nakra
they planted stick-DEF by his grave which NGN
stund ther äfte bar blomstir oc fruit
moment there after bore flower and fruit
(HML 216)

‘They planted the stick by his grave and some moments later the sticks flowered and bore fruit’.
Similarly, NGN can be used alongside a mass noun, meaning roughly ‘some’ or ‘a little’.

(83) thaugik en aff tiänistofolketäpther nakot watn, then went one of servant people after NGN water

tilhwsens behoff, änthathansom

to house-DEF-GEN need yet when he who

watnet hänthenemkomihwset, oc

tillected home came in house-DEF and

sloghtzhthiewethannatak, thawar

poured intothethzwithithannelthenwas

watnet omwänthi blodh

tillected in blood (Linc NR:317-8)

‘Then one of the servants went to fetch some water for the needs of the house, yet when he who carried the water came back to the house and poured it over to another bin, then the water was changed into blood’.

(84) ochtogh aff allom synom storom rikedomom mz

andtook of all her largerriches with

siksalkzenkte, wtan ethklädhemom sit hwwfudh,

selfnothingsbutaclotharoundherhead

oc eth lithetskrinmznakot gull

andEN small box with NGNgold (Linc NR 295)

‘And of all her large riches she took nothing with her but a cloth around her head and a small box with some gold in it’.

NGN therefore seems to be used in a variety of functions found in MS, even though the specific uses, where additional information cannot be provided or is deemed irrelevant are few and interpretable in the light of the modern usage rather than the context itself.

In period I, EN is not yet used in negative polarity contexts, so the competition is mainly between NGN and bare nouns. First in period II EN starts invading the negative polarity contexts, mainly comparisons, and in period III it can be used in all but questions and negation, where either bare noun or NGN are used. Therefore NGN’s domain becomes more limited. Yet, as we have seen in diagram 11, its frequency does not fall; on the contrary, it rises somewhat from period
I to III. One way of explaining this apparent paradox is that in period III both the definite and the indefinite articles are obligatorified in so many functions, and the domain of bare noun so restricted, that some marking is felt to be necessary.

5.9. Anomalous uses of EN: EN defNP

The grammaticalization of -IN into definite article is well underway in period I, with -IN occurring in all textual functions and occasionally with unique referents as well. The new form is spreading through new contexts, some surprising from the MS perspective (see Chapter 4.6 on anomalous uses of -IN). In 4.6.4 we hinted that one of the surprises, the cooccurrence with the numeral EN, will be studied in connection with its grammaticalization into indefinite article.

Let us first consider the following example.

(85) ær han sandær at sak böte atta örtoghaer
is he true to case pay eight örtog
saksøke enum firir vndiruþ øk sax øræ firir
plaintiff EN for tree and six öre for
aldinuid
oaktree

‘If he is guilty of the crime he shall pay eight örtog
for the tree without fruit and six öre for the oaktree to one plaintiff’.

Collin and Schlyter edition of AVL has here saksøke enum (Collin and Schlyter 1827:61). A double-check with the facsimile edition of the manuscript confirms that to be the correct reading.

A specific reading of the NP is difficult (= a plaintiff, if specific, why not definite, the plaintiff?), a numeral is more likely (= one plaintiff), though there is no contrast here (e.g. two charges—one plaintiff). More surprisingly, in a similar example in YVL, both -IN and EN are to be found within one NP:

(86) Falder klocka rummi nither i houoth manni böte sokn
falls bell room down in head man pay parish
fore III marker arfuønom enum øn han dór
for three marks heir-DAT EN if he dies

About 1/3 of öre.
Anomalous uses of EN: EN defNP

aff
of

(YVL KB:18)

‘If a church bell falls down on a man’s head, the parish shall pay three marks to the heir if the man dies of it’.

The combination EN + defNP is not foreign to MS, however it is only possible with the weak form of EN, namely ena ‘the one’, as in the following example:

(87) Trots att han har brutit ena benet..
although that he has broken one-W leg-DEF

‘Although he has broken one leg (he still managed to finish the race)’.

Larm (1936) notes one such example in classical OS from a relatively late text, Kristoffers landslag (ca 1440):

(88) ena syskionebarnet
one-DEF sibling-child-DEF

(after Larm 1936:43)

It seems that the classical OS (period I) system allowed the combination of strong form of EN and -IN where MS only allows the weak form (compare notes on strong adjectives by definite nouns in 4.6.2). The meaning is partitive, similar to ‘one of’ or possibly ‘the first’, as it is often followed by further enumeration, see the following example.

(89) Alt þæt guþ gaf þøm at ægha skipþo þe þrem lutum ok skipþo: ena delen
tel guz monstar ok þienist annan
to god-GEN monastery and service other
fatoko folke þripia sik ok sinom
poor-DAT people-DAT third self and own-DAT
boskap
livestock

(Bur 3)

‘All that God gave them they divided into three parts and shared: one part with God’s monasteries and services, the second with poor people and the third with themselves and their livestock’.
In (86), where no further heirs are mentioned, the intention may have been to point out that the fine was due to one heir only, even if there were more than one potential heirs meaning, more or less, ‘one of the heirs’. Such a structure is impossible in MS, but is documented both in ON (examples (90-94) and OS (example, (95); all examples collected in Perridon 1989:197):

(90) Hundarnir .. töku hvern fuglinn
dogs-DEF took each bird-DEF
‘The dogs took each of the birds’.

(Heimskr. II 164, cf Lundeby 1965:105)

(91) Hoðr heitir einn ásinn
Hoðr is called one Æsir-DEF
‘One of the Æsir called was Hoðr’.

(Sn Edda 17, 32, cf. Nygaard 1905:32)

(92) hann sendi aptr suman mjoðinn
he sent after some mead-DEF
‘He sent for some of the mead’

(Sn Edda 73, 29, cf. Lundeby 1965:116)

(93) drap eg þá marga vampana
killed I then many wolves-DEF
‘Then I killed many of the wolves’.

(Njála 139, 27, cf. Lundeby 1965:117)

(94) Engi knút fekk hann leyst ok engi
not-one knot got he loosened and not-one

álarendann hreyft
leather-strap-DEF moved
‘He could not untie any knot nor move any leather strap’.

(Sn Edda 29, 31, cf. Nygaard 1905:32)

(95) Fiärdha är att ängin ouinnin ingange ginom
fourth is that not-one enemy-DEF entered by

portana
gates-DEF

‘The fourth is that none of the enemies may enter through the gates’

(Birgitta III, 240, cf. Ågren 1912:54)
5.10. Conclusions

A detailed study of the functional spread of EN in OS shows that the model of grammaticalization of the indefinite article from a numeral as proposed by Heine 1997 can be applied to the OS material. The first non-numeral uses of EN are indeed as a presentation marker which is used to establish discourse referents with a long lifespan, measured by the average number of further mentions of the discourse referents. The EN-phrases found in this function are mainly sentence- and text-initial subjects. Gradually, the average number of further mentions decreases, independent of the grammatical function of the EN-phrase. EN-phrases are used to establish new discourse referents with a short lifespan (non-specific discourse referents).

The stages of grammaticalization of the indefinite article in Old Swedish are given in Table 35 below.

Table 35. Stages of grammaticalization of EN in periods I-III

<table>
<thead>
<tr>
<th>Stages of grammaticalization of the indefinite article</th>
<th>Functions</th>
<th>Spread to context</th>
<th>Obligatorification in context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Period I</td>
<td>Period II</td>
</tr>
<tr>
<td>Stage II</td>
<td>Presentative marker</td>
<td>(+)</td>
<td>+</td>
</tr>
<tr>
<td>Stage III</td>
<td>Specific marker</td>
<td>(+)</td>
<td>+</td>
</tr>
<tr>
<td>Stage IV</td>
<td>Nonspecific marker</td>
<td>–</td>
<td>(+)</td>
</tr>
<tr>
<td>Stage V</td>
<td>Generalized article</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

We have at several points made references to the stage of grammaticalization of the definite article when discussing the development of the indefinite one. The interactions between the two grammaticalizations will be discussed in 6.4.
6 Conclusions

6.1. Introduction

Articles are relatively late additions to the grammars of Indo-European languages. In fact, their grammaticalizations may almost always be studied, at least in part, in the text heritage of a given language. Although the Indo-European languages are inter-related, the category of definiteness is not part of their common heritage but is an innovation in each and every one of them. We may either choose to consider this innovation as an inevitable part of some other major development—the ‘drift’ towards a particular type; or we may look upon it as an areal rather than genetic feature (Dahl 2007:34).

In fact, there is a relatively neat diachronic progression in the appearance of definite articles from the Eastern Mediterranean to north-western Europe, basically in the order Semitic → Greek → Romance → Germanic, suggesting a rather slow expansion wave which took about two thousand years to complete. (Dahl 2007:34).

Strikingly, the languages where definite articles grammaticalize are concentrated in Western Europe; the suffixed articles only appear peripherally, in Scandinavia and the Balkans whereas in the rest of the territory the definite articles are preposed.

Among Indo-European languages that have developed the definite article, most have also developed the indefinite article. The exceptions include Bulgarian and Icelandic. Cross-linguistically, languages with only definite articles are more common than those with both definite and indefinite articles, and languages with only indefinite articles are even more rare (see Chapter 1, WALS).

The present analysis is based on Swedish texts and spans from 1225 to 1529 AD, the so-called Old Swedish period (fornsvenska). For the purpose of the study, OS has been subdivided into three periods, period I (1225-1375), period II (1375-1450) and period III (1450-1526), based on earlier practice and tradition. The subdivision has revealed differences between the three periods, with period II being
the time of the most dramatic and radical change in the use of both -IN (the grammaticalizing definite article) and EN (the grammaticalizing indefinite article).

It has been shown in chapters 4 and 5 that the grammaticalizations of the definite and the indefinite articles in Old Swedish largely follow the assumptions made for these two developments, including the sources and the stages, as identified in the works of Greenberg 1978, Givón 1981, Lehmann 1995 and Heine 1997. The source of the definite article is the distal demonstrative (h)inn and the source of the indefinite the numeral en ‘one’. The development proceeds through the stages given in Chapter 2. A close inspection of the sources reveals more details of both developments.

As the statistics show, languages with both definite and indefinite articles dominate among article languages (194 languages according to WALS); there are many languages with only definite article (98); but relatively few languages with just the indefinite article (45). It seems that the indefinite article grammaticalizes when the grammaticalization of the definite article has already started and that the two articles’ developments are interconnected. In this final chapter we take a closer look at the interdependencies between the two grammaticalizations.

The chapter is organized as follows: in 6.2 we summarize the most important results concerning the grammaticalization of the definite and indefinite article presented in chapters 4 and 5. In 6.3 the models of grammaticalization of both definite and indefinite article are discussed. 6.4 presents the interplay between the two grammaticalizations and 6.5 the connection with other grammatical changes that take place in the OS period. Finally, in 6.6 the long-term perspective and possible future research are discussed.

6.2. Overview of the grammaticalization of the definite and indefinite articles in Old Swedish

6.2.1. Summary of the developments

The grammaticalization of the definite article seems to have started with the extension of the deixis of the pronoun from purely situational to textual and the marking of anaphora. This earliest development predates the oldest sources available (see Chapter 3) and can only be reconstructed. What is seen in the oldest texts, however, is not a clear-cut picture. It seems that -IN is more likely to appear in contexts where the referent, though introduced earlier in the text, is not easily accessible, with accessibility measured by the number of intervening syntagms. The role of -IN thus seems to have been the ‘activation’ of a referent present in
the discourse but not immediately available. The higher the accessibility, the more competition between -IN and the demonstrative sá (den). The growing dominance of -IN in direct anaphoric contexts coincides with the obligatorification of their marking and can be said to be completed first in Period II.

As early as in Period I -IN is used in contexts here termed indirect anaphora. By period III their marking is obligatory. The development is largely dependent on the ‘anchor’ of the reference which triggers the definite. The first instances of indirect anaphora involve part-whole relationships, with the exception of body parts (where the marking is either zero or a possessive pronoun), action-instrument relationships, etc.

Both direct and indirect anaphora constitute textual uses of -IN. Both become obligatory in period II. Non-textual uses, first only by unique referents, though found in period I, first become obligatory in period III. Referents whose uniqueness is more local (such as ‘the district’) are marked earlier (period I), whereas those whose uniqueness is more universal (‘the king’) remain largely unmarked in period I.

EN is largely employed as numeral ‘one’ (legal prose) and as presentation marker (religious prose) in period I. In religious texts it appears occasionally as a specific marker, but no instances of non-specific use are noted. In period II both the presentation and the specific function are obligatory and the first instances of nonspecific uses of EN are found.

In MS, generics can be marked in five different ways (bare singular, bare plural, definite singular, definite plural, indefinite singular). Such a complex system is not yet found in OS. However, as early as in period II we find the first instances of -IN used generically. In period III, the first occurrences of EN in generic-like contexts have been noted, though these may be more properly classified as ‘critical contexts’, where the generic interpretation is possible, but not necessarily intended.

### 6.2.2. Textual origins of the articles

Both the definite and the indefinite article originate as textual devices. The first function adopted by the demonstrative at the onset of the grammaticalization is the marking of referents familiar from the discourse (direct anaphora). The function is to bind the text together, making it coherent. In its role of a presentation or specificity marker, EN also seems to be a textual device which marks the discourse referents that are going to be prominent in the discourse.

Surprisingly, there is some competition between EN and -IN in the presentative role, though it is marginal and limited to one genre (legal prose), one period
Overview of the grammaticalization of the definite and indefinite articles in Old Swedish

(Period I) and a handful of lexemes (the most prominent among them the noun *bondi* 'yeoman'). What we often see is that high values for -IN coincide with high values for EN. This may be taken to be a hallmark of a highly narrative text, where many referents are introduced and followed up. It is particularly clear in texts like Jart (a religious prosaic text from period II), which is a collection of short narratives with a great number of referents (usually at least three per story). In texts of lower narrativity we find correspondingly lower values for textual -IN (though they may be high for non-textual ones, see religious prose, in particular Linc JB from period III).

With time, as both definite and indefinite articles spread outside the textual domain, the character of the text does not influence their frequencies that much; obligatorification of the definite and indefinite marking means that the genre, which may still play a role in the percentage of a particular type of definiteness, will not be enough to strongly reduce the frequencies. In other words, the results may differ because though the text introduces a number of discourse referents who are mentioned frequently, it does not discuss philosophical or theological questions. The values will be high for textual definites but if a non-textual definite occurs it will necessarily be marked. On the contrary, a philosophical dispute may contain a low number of textual definites, though these will likewise be necessarily marked.

### 6.2.3. Stages of the grammaticalization of the articles in OS

In Chapter 2 we have identified stages in the development of the definite and indefinite articles and applied them in Chapter 4 and 5. The results are collected in tables 31 in chapter 4 and 35 in chapter 5. These results confirm yet again that the main development in the grammaticalization of -IN into a definite article occurs in periods I and II and that period III is a time of stabilisation of the form. Interestingly, 1375-1450 seems to be the time when many other major changes are resolved (see 6.5): the case system dismantled, the word order stabilized. A detailed study of how these changes interplay falls beyond the scope of this book, but it would no doubt be a worthwhile project to study them against each other and not in separation. It is to be noted here that the subdivision of *yngre fornsvenska* (1375-1526) into period II and III, which has become a more or less common practice, helps to better delimit the extensive changes in OS. By separating the earlier and later parts of the younger OS, we were able to see more clearly how the definite article undergoes a most dramatic grammaticalization, gaining ground in practically all its modern functions, and how its use within these functions stabilizes in period III. The grammaticalization of the indefinite article, on the other hand, gains momentum in period II but peaks first in period III. Study of its
stabilization period falls outside the scope of this study. From other studies it may be seen that what is here termed period II is the time of some tumult in language history, with many changes taking place at the same time.\footnote{The present study focuses on the linguistic changes. It should be noted that they are highly dependent on the social changes. One of the most dramatic social changes in Scandinavia and Europe of that time was the plague (the Black Death), which devastated the population between 1348 and 1350. 1375, the beginning of the younger Old Swedish period, is the time when the first generation of people who were born after the plague, or who have survived it as young orphans and were probably raised by people other than their families, reached adulthood. Their language could be dramatically different from the language of their parents’ generation.}

### 6.2.4. Note on the co-textual definiteness

Throughout the study we have used the concept of non-textual definiteness to refer to the non-anaphoric functions of -IN (direct or indirect). In discussing the use of -IN in religious prose, in particular in period II, we have also used the term co-textual to describe the uses of -IN that were not anaphoric and yet seemed to be dependent on the text and its purpose.

The so-called co-textual definites show how intricately the grammaticalization of the articles is interwoven with texts. These definites are found even where there are no linguistic signals in the text. The text itself is however a signal on how the definites are to be interpreted. Naturally we know little about the situation in which the text was written, for whom and in what circumstances it was read, meaning that many definites that could be understood as allusions to something known to every reader at the time the text was written, may be either opaque to us or could also be misinterpreted with to our scanty knowledge of the times. Yet we are able to see that the interpretation of these definites is dependent on the text, even if it is not linguistic.

### 6.3. Discussion of the models of grammaticalization

The models of grammaticalization of articles quoted in this study are Greenberg’s 1978 for the definite article, and Givón 1981 and Heine’s 1997 for the indefinite article.

(1) \textit{demonstrative} \textgreater \textit{definite article} \textgreater \textit{specific article} \textgreater \textit{noun marker} \hfill \textit{(after Greenberg 1978)}
to a specific article, via the definite article. The model does not specify the interim stages of grammaticalization between a demonstrative and an article.

Lyons 1975 sees the first stage of grammaticalization in the use of the demonstrative as an anaphoric marker. In his account, the text is treated as a situation within which the speaker (writer) helps the hearer (reader) to navigate. This prediction is verified by the OS data, where it can be seen that the textual uses (both direct and indirect anaphora) dominate heavily in period I and non-textual uses appear first in periods II and III.

However, -IN is found in functions other than textual as early as in period I. This may be due to the fact that at the time of writing the first texts the grammaticalization was already advanced and -IN had spread beyond its original domain.

As seen previously, in the oldest texts (legal prose), -IN can be used in contexts where MS would prefer the indefinite article. Although this use is mainly found with the weak noun bondi ‘yeoman’ and may be explained away by the wish to make a case form clear, it is frequent enough to be considered. In these contexts -IN seems to function as a specific rather than definite article which runs counter to Greenberg’s model, which aims mainly to describe the final stages of the grammaticalization of the definite article and their new role as gender markers (as the title of the paper suggests). The development is studied mainly for Niger-Congo languages; and its early stages—from demonstrative to a definite article and from a definite article to a specific one, are largely reconstructions. Its application on the OS material is limited at best (see 6.4 for further discussion).

The model of the grammaticalization of the indefinite article as proposed by Heine 1997 is a more detailed version of that in Givón 1981.

As we have noted in Chapters 2 and 5, more detailed criteria, applicable in diachronic studies, are necessary to differentiate between the presentative and specific functions on the one hand and the specific and non-specific on the other. To do this, this study uses Heim’s notion of ‘lifespan’ of a discourse referent (Heim 1982). The study of the number of further mentions of a discourse referent established by an EN-phrase reveals that in its earliest instances, EN established discourse referents of the longest lifespan, but by period III the average lifespan...
had fallen from over 6 to 2. In this respect, Heine’s model of the grammaticalization of the indefinite article is confirmed in the OS material.

Yet for the OS data it must be noted that EN can be regarded as a presentative marker only in connection with the grammatical function of the EN-phrase (subject) and its position in the sentence and text (initial). EN-phrases which are not subjects and/or are not placed sentence-initially do not establish new discourse referents with a long lifespan.

Both Lyons and Heine place the origin of the articles in the textual domain: as markers of discourse referents already present in the discourse (anaphora) or to be discussed (presentation). The latter marker applies to prominent and persistent discourse referents and it is possible to hypothesise that the former did as well during the onset of its grammaticalization. A reasoning along these lines can be found in Johnsen’s study of the origins of the definite article in OI (Johnsen 1976) as well as in Leiss’ study of definiteness in the same language (Leiss 2000). Johnsen implies that -IN was originally used to emphasise prominent discourse referents already mentioned. Whether that was the case in OS is impossible to tell on the basis of the material available. The grammaticalization of the definite article as witnessed by OS period I texts is already advanced and it is difficult to see that prominent discourse referents are preferred and marked by -IN more often than the less prominent ones. On the other hand, the grammaticalization of the indefinite article originating with the discourse referents ‘prominent-to-be’ makes this hypothesis very plausible. Both developments could then be seen as originating in the need to single out important discourse referents, which is a pragmatic factor.

6.4. Interplay between the grammaticalizations

In OS the grammaticalization of the indefinite article starts when there is already an embryonic definite article. Similar time difference is found in other Indo-European languages. It would be of great interest to see whether some common turning point in the development of the definite may lay the ground for the development of the indefinite. To my knowledge, there are no such studies yet. From the present study we may note that the grammaticalizations of both definite and indefinite articles originate within the textual domain and that the numeral expands into presentation contexts as -IN stabilizes in its textual functions. There is even, as mentioned above, a certain amount of functional overlap between these two forms (note also how English uses the demonstrative this to introduce new, salient discourse referents, showing that demonstratives—and -IN is, to begin with, a demonstrative—may serve as presentation markers). It is possible to hypothesise that this functional overlap between -IN and EN leads to an attempt at giving the two forms different interpretations, holding them apart and creating
rules for their use. In this attempt notions of given (in the discourse) vs. new (in
the discourse) are employed, later extended to the wider notions of known (from
the discourse or otherwise) vs unknown.

Even if the functional overlap between -IN and EN were minimal, the two
grammaticalizations in Swedish seem to be dependent on each other. The primary
function of the demonstrative that activates the grammaticalization chain leading
to the definite article is to mark referents given in the discourse. The primary func-
tion of the numeral that initiates its grammaticalization into the indefinite article is
to mark those referents that are new and topical. Both items, -IN marking anapho-
ra and EN functioning as marker of cataphora, thus serve to create coherence and
continuity in text: the presentative function more or less means that the discourse
referent established by the EN-phrase will be picked up in later discourse (there is
again a clear parallel to the development of English demonstrative this).

Although the overlaps between -IN and EN, which are unfamiliar in the MS
article system, may seem incomprehensible to us, it is important to remember that
the articles do not grammaticalize in a void—they appear in a system which is al-
ready capable of expressing any meaning (as the uniformitarian principle claims).
Therefore the development takes place not because something is wanting; rather,
the new form must establish itself against other, older ones. Therefore in the old-
est texts we see high ratio of demonstratives, which are in competition with -IN,
whereas in period II the conflict is resolved and the two appear within one NP
and the conflict we witness is between -IN and possessive pronouns, which also
show exceptionally high figures. The latter conflict is resolved in period III where
a functional division is reached (e.g. by inalienable possession) and the conflict
within the NP resolved so that the appearance of one within the NP excludes the
other (apart from dialectal NPs with postposed possessive).

If Greenberg’s model is correct and the grammaticalization goes through
definite to specific article, development beyond the definite is checked in Swed-

ish. Among the inhibitive factors are undoubtedly the further grammaticalization
of the indefinite article—the presence of the indefinite narrows the scope of use
of the definite article. Additionally, we note that the growth in literacy may also
have played a part in freezing the development, especially in the standard lan-
guage.

Yet there are Swedish vernaculars where as early as 1878 an overuse of defi-

nite article has been noted in contexts where the standard variety would prefer the
indefinite one or a bare noun, e.g.

(3) Kva ha et tjøft i stádin? Åttren o grýnen.
What have you bought in town-DEF pea-DEF.PL and grain-DEF.PL
(Närpes dialect, Freudenthal 1878:137, quoted in Dahl 2007:25)
It seems that in the non-standard, peripheral varieties the spread of -IN to new functions has continued to include even the contexts reserved for the indefinite article. This spread of the definite article beyond its purely ‘definite’ contexts is noted by several authors writing on Swedish dialects, among others Levander (1909, Elfdalian dialect), Hummelstedt (1934, Ostrobothnian dialect), Delsing (1993), Nicula (1997, Näpnes dialect), Dahl 2007. Delsing terms this use a ‘partitive’ one (1993:50), similar to the French use of the definite article. In later work he notes that this article use is found with uncountable nouns, plurals and singulars that denote un-delimited or arbitrary quantities (Delsing 2003:15).

What we find in Scandinavian vernaculars, however, is an expansion of the range of uses of definite articles that goes in a slightly different direction and cannot be described in terms of “specificity” in any sense. The Scandinavian development therefore is of considerable interest for our understanding of the role of definite articles in grammaticalization processes. (Dahl 2007:32)

The definite and the indefinite articles are, as their names suggest, often viewed as denoting mutually exclusive grammatical functions (Heine 1997:70), yet they can occur in some contexts as functionally equivalent or near-equivalent items. Heine 1997 shows it with respect to the final stages of grammaticalization of both, where there is a certain overlap in the generic functions (though in MS there are still different notions attached to either the definite or the indefinite in this context). The present study has shown that the earliest stages of grammaticalization of the definite may also show a similar tendency, where the definite-to-be is used as the presentative marker, in competition with the numeral ‘one’ and other items.

### 6.5. Interplay with other grammatical changes

In periods I-III the Swedish language undergoes a series of changes involving both syntax and morphology. The case system is lost (Norde 1997, Delsing 2002, Skrzypek 2005), word order fixed as VO (Delsing 1999) and the subject becomes obligatory (Håkansson 2008). The rise of the definite article has often been placed in the context of the demise of case morphology: in the European languages the two categories are found only in a handful. Yet no studies can prove conclusively that the two processes—the rise of the articles and the decline of case, are indeed dependent on each other. If that were the case, some functional overlap between the categories would be necessary. This functional overlap may be the textual role of both: the case binds the text together through agreement and definiteness by the marking of given and new information. They may exist side by side, as in German
or Icelandic, but one of them may also be enough. Note though that the cohesion created by articles goes beyond the clause or sentence and is a discourse device, whereas case agreement operates only within clauses/sentences. Thus definiteness is a category that, though not as precise as case on the clausal level, is useful on a level which case agreement does not reach.

6.6. The long-term perspective

The study covers only the Old Swedish period. By the early 16th century an article system emerges, whose major properties seem very similar to the MS one. And yet the frequencies of the articles, especially the indefinite article, are significantly lower than the corresponding MS ones. Is the system emerging in the early 16th century already identical with the MS one?

For the definite article, despite the lower frequencies, its uses are identical in period III and MS. The factor that may be important is the structure of the defNP and the co-occurrence of different definite markers (see Appendix for an analysis of the evolution of the NP).

For the indefinite article, it is not yet fully obligatory in non-specific contexts in period III and is still in variation with bare noun there. A more detailed study of the negative polarity contexts may render a more detailed picture of further development.

Finally, generics seem to be mainly expressed by bare nouns through OS, which naturally lowers the frequencies of the articles, though a more detailed study is called for. The MS system, where generics can be either expressed by bare nouns (in the singular or in the plural), or by an article form (definite singular or plural or indefinite), is not yet formed in OS and its formation needs further exploration.

6.7. Concluding remarks

The view of definiteness presented in this book is of necessity textual—the connection between text and form is of primary importance, as the names given to different categories: textual, non-textual, even co-textual, suggest. This dependence is created by the fact that (in)definiteness in historical studies is only accessible through written texts, and cannot be studied ‘dynamically’—with native speakers giving their views and completing questionnaires. Since we have previously stated that the origins of the definite article lie beyond the textual heritage, we must conclude that what we have studied here has been the development of the article in its textual domain. The development in the spoken language could have
been quite divergent and filtered through to the written one only afterwards—maybe the ‘somewhat late’ appearances of definite uniques were in fact not late at all.

The rise of definiteness provides a highly interesting research topic because it is a relatively new phenomenon in languages and can at least partly be studied in the texts, not just reconstructed. Even in the Scandinavian languages, where, as we have seen, the cliticization of the demonstrative predates the oldest sources, its functional development can be studied in the oldest texts. The indefinite article is even younger (as we have noted before, its rise as a presentative marker is dependent on the presence of a textual marker of ‘givenness’ in the language). This shows that the articles are connected with texts—not just sentences, but texts, where longer stories with many heroes are told. The rise of the articles seems also to be connected to literacy and production of texts. It is no coincidence that the texts which show the earliest examples of article use are narratives (religious or profane) and not laws.

The study presented here is by no means exhaustive and while it hopefully answers some questions concerning the grammaticalization of the definite and indefinite articles, at the same time it raises probably just as many. In this final part I have outlined the major issues for future research, concerning the Swedish language as well as other article languages. A closer inspection may reveal more facts about the cross-linguistic similarities in the development.
Appendix: Notes on NP types in OS

NP types in period I

In contrast to the MS NP with a fixed word order, the OS NP shows much more variation in respect to both the order of the elements and their co-occurrence. This variation gradually wears off, however in period I there is still a greater preference for postposition of determiners than in later periods (where postposition gradually disappears) and a greater freedom in how the different determiners and modifiers can be combined. The types of NPs found in period I texts are collected in Table 36 below. In the first column, the MS NP types are given, in the second those found in period I and examples of the latter in the third column. The OS NP types no longer productive in MS are given in bold.

Table 36. Types of NPs in period I

<table>
<thead>
<tr>
<th>MS NP types</th>
<th>Period I NP types</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>N</td>
<td>the same</td>
<td></td>
</tr>
<tr>
<td>N-en</td>
<td>the same</td>
<td></td>
</tr>
<tr>
<td>en (ADJ) N</td>
<td>en (ADJ) N</td>
<td>en syndogha kona (Bur 133)</td>
</tr>
<tr>
<td></td>
<td>N en</td>
<td>pyno ena (Pent 179)</td>
</tr>
<tr>
<td></td>
<td>en (ADJ) N-en</td>
<td>een delen (Bur 3)</td>
</tr>
<tr>
<td></td>
<td>N-en en</td>
<td>arfuanom enom (YVL KB:18)</td>
</tr>
<tr>
<td>ADJ N</td>
<td>ADJ N</td>
<td>vatnløsan brwn (Pent 178)</td>
</tr>
<tr>
<td></td>
<td>ADJ-a N(^1)</td>
<td>sota frukt (Bur 140)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>suartæ slagh (ÄVL BB:5)</td>
</tr>
<tr>
<td>ADJ-a N-en</td>
<td>ADJ N-en</td>
<td>østir riken (Pent 159)</td>
</tr>
<tr>
<td></td>
<td>ADJJa N-en</td>
<td>myrke husit (Pent 184)</td>
</tr>
</tbody>
</table>

\(^1\) This type is marginally found in MS, limited to interjections only, e.g. *Gamle man!* ‘Oh, old man!’
<table>
<thead>
<tr>
<th></th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DET N-en</td>
<td>den N&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>poss N</td>
<td>N poss</td>
</tr>
<tr>
<td>poss ADJ-a N</td>
<td>poss ADJ-a N</td>
</tr>
<tr>
<td>(DET) samma (ADJ-a) N</td>
<td>(DET) samma (ADJ-a) N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DET N-en</td>
<td>N den</td>
</tr>
<tr>
<td>poss N</td>
<td>N-en</td>
</tr>
<tr>
<td>poss ADJ-a N</td>
<td>N-en</td>
</tr>
<tr>
<td>(DET) samma (ADJ-a) N</td>
<td>N-en</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DET N-en</td>
<td>þæn tima (Bur 149)</td>
</tr>
<tr>
<td>poss N</td>
<td>sak þerre (ÄVL FB:9)</td>
</tr>
<tr>
<td>poss ADJ-a N</td>
<td>hit malit (OgL RB:18)&lt;sup&gt;3&lt;/sup&gt;</td>
</tr>
<tr>
<td>(DET) samma (ADJ-a) N</td>
<td>þæn fiken (Bur 153)</td>
</tr>
<tr>
<td>DET ADJ-a N-en</td>
<td>hin ADJ-a N</td>
</tr>
<tr>
<td>poss N</td>
<td>þæt hælgha cors (Bur 140)</td>
</tr>
<tr>
<td>poss ADJ-a N</td>
<td>þæn hælgha mann (Bur 188)</td>
</tr>
<tr>
<td>(DET) samma (ADJ-a) N</td>
<td>þætta bilætte (Bur 189)</td>
</tr>
<tr>
<td>poss ADJ-a N</td>
<td>thæne bruwn (Pent 162)</td>
</tr>
</tbody>
</table>

If the dominant pattern in Runic Swedish was postposition of both adjectives and pronouns in the NP, period I is the time of change from the original postposition of elements in the NP towards their preposition, which can be seen from the fact that both types are richly represented in the texts. Furthermore, the three exponents of definiteness, the weak adjectival declension, the suffix -IN and the definite determiner, which today all co-occur

---

<sup>2</sup> This type is only found in MS before a relative clause *den man som kom imorse* ‘that man who came in the morning’.

<sup>3</sup> Isolated occurrence, the only one in the material (whole texts of OgL, UL and DL searched)

<sup>4</sup> This type is retained in some lexicalized phrases, most connected with religion, e.g. *den heliga skrift* ‘the holy book’ = the Bible.

<sup>5</sup> Although the standard language only allows N after *denna* in MS, there are many dialects where the definite form is used (Hirvonen 1987).

<sup>6</sup> Possible in dialects.
within one NP, can be used exclusively to mark the NP in period I. Thus the type *den N*, in MS only possible before a relative clause and *ADJ-a N*, today limited to exclamations, is possible in period I.

**NP types in period II**

In period II, the tendency to preposition is stronger than in period I, and fewer combinations are possible.

<table>
<thead>
<tr>
<th>MS NP types</th>
<th>Period II NP types</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>N</em></td>
<td>the same</td>
<td>kejsaren (KM)</td>
</tr>
<tr>
<td><em>N-en</em></td>
<td>the same</td>
<td>haffwith (KM)</td>
</tr>
<tr>
<td><em>en (ADJ) N</em></td>
<td><em>en (ADJ) N</em></td>
<td>en storan wall oc wäl kläddan (KM 251)</td>
</tr>
<tr>
<td><em>ADJ N</em></td>
<td><em>ADJ N</em></td>
<td>ärliker man (SVM)</td>
</tr>
<tr>
<td></td>
<td><em>ADJ-a N</em></td>
<td>faghra kăldo (HML 193)</td>
</tr>
<tr>
<td></td>
<td><em>ADJ-a N-en</em></td>
<td>halte mannén (HML 191)</td>
</tr>
<tr>
<td><em>DET N-en</em></td>
<td><em>N-en den</em></td>
<td>the kirkio (KM 251)</td>
</tr>
<tr>
<td></td>
<td><em>den ADJ-a N</em></td>
<td>thán kărlek som (Jart 13)</td>
</tr>
<tr>
<td></td>
<td><em>den ADJ-a N-en</em></td>
<td>disk then som (KM 252)</td>
</tr>
<tr>
<td></td>
<td><em>hin ADJ-a N-en</em></td>
<td>thy yflätino (Jart 4)</td>
</tr>
<tr>
<td><em>denna (ADJ-a) N/N-en</em></td>
<td><em>denna (ADJ-a) N</em></td>
<td>thässa kirkio (KM 251)</td>
</tr>
<tr>
<td></td>
<td><em>denna (ADJ-a) N-en</em></td>
<td>thässa flodhena (Jart 12)</td>
</tr>
<tr>
<td><em>poss N</em></td>
<td><em>poss N</em></td>
<td>sit hofwdh (KM 252)</td>
</tr>
<tr>
<td></td>
<td><em>(DET) samma (ADJ-a) N-en</em></td>
<td>thz sama aarith (KM 250)</td>
</tr>
</tbody>
</table>

**NP types in period III**

We have seen that the great variation in the form of the NP, characteristic of period I, was greatly reduced in period II. Period III continues in the same vein: all determiners, attributes and modifiers are preposed. The exceptions include NPs with more than two adjective modifiers, which are then split, the first occurring prenominally and the others postnominally.
For nordan fiäll ligger et slot som sägward
for northern mountain lies a castle which

heter thz atte hin rikä brynild hin fagra
is-called it owned DET rich Brynhild DET beautiful

oc hin wiisa
and DET wise

(Did 16)

‘At northern mountain there lies a castle, called S, owned by the rich, beautiful and wise Brynhild’.

Although NP still shows some variation in period III, the type *den gamle mannen*, which is the unmarked one in MS, starts to dominate. In his detailed study of the variation between *den gamle man*, *den gamle mannen*, *gamle mannen* Hirvonen (1997) gives the following numbers for their frequencies in all three periods (table 38).

Table 38. Variation *den gamle man*, *den gamle mannen*, *gamle mannen* in OS, after Hirvonen 1997

<table>
<thead>
<tr>
<th>Period</th>
<th><em>den gamle man</em></th>
<th><em>den gamle mannen</em></th>
<th><em>gamle mannen</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Period I(^7)</td>
<td>54%</td>
<td>14%</td>
<td>32%</td>
</tr>
<tr>
<td>profane prose</td>
<td>63%</td>
<td>28%</td>
<td>9%</td>
</tr>
<tr>
<td>religious prose</td>
<td>22%</td>
<td>37%</td>
<td>41%</td>
</tr>
<tr>
<td>Period II</td>
<td>21%</td>
<td>42%</td>
<td>37%</td>
</tr>
</tbody>
</table>

As can be seen from Table 38, the MS unmarked type *den gamle mann* gradually increases in frequency from period I to period III, from least to most frequent, in Hirvonen 1997 to the most frequent. Hirvonen’s results also show that the religious literature is more ‘modern’ even in this respect, since the *den gamle mann* type occurs twice as frequently as in the legal prose.

\(^7\) Although the standard language only allows N after *denna* in MS, there are many dialects where the definite form is used (Hirvonen 1987).

\(^8\) Hirvonen bases his study of period I on results presented in Larm 1936, and he divides the material into profane, religious and poetic. Profane prose includes legal texts and *Konungastyrelse*.
Sources


All source texts apart from JB available also from Fornsvenska textbanken, www.lu.se/nordlund.

Modern Swedish short stories
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


Johnsen, Gail Perkins. 1976. *The development of the definite article in Old Norse.* Diss. The University of Rochester.


REFERENCES


List of tables and list of diagrams

List of tables

Chapter 1
Table 1. Distribution of definite and indefinite articles in the world’s languages
Table 2. Excerpted texts
Table 3a. Search strings for KWIC concordance creator
Table 3b. Paradigm of *en ‘one’ in Old Swedish

Chapter 2
Table 4. Stages and contexts in the grammaticalization of the definite article

Chapter 3
Table 5a. The inflectional paradigm of *(h)inn
Table 5b. The inflectional paradigm of *sá
Table 5c. The inflectional paradigm of the demonstrative *sjá (*pænne)
Table 6. The inflectional paradigm of enn ‘one’
Table 7. The definite nouns in Old Swedish
Table 8. Proposed dating of the cliticization of *(h)inn

Chapter 4
Table 9. Occurrences of -IN in periods I-III
Table 10. Occurrences of -IN throughout the texts in Periods I-III
Table 11. The distribution of -IN in AVL by number, gender and case
Table 12. Direct anaphora in KM
Table 13. The marking of direct anaphora in Linc NR
Table 14. Anaphoric marking in Troj
Table 15. The marking of the noun *arvi / arvinge ‘heir’ in period I, legal texts
Table 16. Indirect anaphora in period I, some trigger types
Table 17. Triggers in indirect anaphora in period II
Table 18. Distribution of -IN and -Ø with unique referents in Bur
Table 19. Unique referents in UL and OgL after Larm 1936
Table 20. Unique referents in UL, OgL and DL
Table 21. Distribution of -IN and -Ø with unique referents in religious prose, periods I-III.
Table 22. Distribution of the noun *bondi with and without –IN in DL, UL and OgL
Table 23. The inflectional paradigm of *bonde ‘yeoman’
Table 24. Noun *bondi in the nominative in all OS legal texts
Table 25. Noun *bondi* in oblique cases in all OS legal texts
Table 26. Marked and unmarked forms of the noun *klokkare* ‘sacristan’ in OS legal texts
Table 27. Marked and unmarked forms of the noun domare ‘judge’ in OS legal texts
Table 28. Distribution of -IN in textual and non-textual functions in OS
Table 29. Singular and plural -IN phrases in OS
Table 30. Grammatical functions of -IN phrases in OS
Table 31. Stages of grammaticalization of -IN in periods I-III

**Chapter 5**
Table 32. Occurrences of EN (adnominal) in Period III.
Table 33. Occurrences of EN throughout the texts in Periods I-III
Table 34. Comparison contexts in period II
Table 35. Stages of grammaticalization of EN in periods I-III

**Appendix: Notes on types of NPs in OS**
Table 36. Types of NPs in period I
Table 37. Types of NPs in period II
Table 38. Variation *den gamle man, den gamle mannen, gamle mannen* in OS, after Hirvonen 1997

**List of diagrams**

**Chapter 1**
Diagram 1. Periodisation of the North Germanic languages before 800 AD.
Diagram 2. Periodisation of the Swedish language

**Chapter 2**
Diagram 3. Model of the grammaticalization of the indefinite article according to Heine
Diagram 4. Model of the grammaticalization of the indefinite article according to Givón

**Chapter 4**
Diagram 5. Noun *bondi* in OS legal prose, oblique cases

**Chapter 5**
Diagram 6. Grammatical functions of EN-phrases in OS, periods I-III
Diagram 7. Grammatical functions of EN-phrases in OS texts
Diagram 8. Initial and non-initial EN-subjects in periods I-III
Diagram 9. Average number of further mentions of a referents introduced by EN-subject in OS
Diagram 10. Average number of further mentions of a discourse referent introduced by an EN-phrase, periods I-III
Diagram 11. Frequencies of adnominal EN and NGN in OS
Index

A

abstract 149, 172
accessibility 27, 48, 50, 82, 84, 92-94, 100, 101, 114, 167, 196, 197
actualization 82, 83, 90, 103
adnominal 22, 70, 71, 156-158, 185
agentive 97
amorphous 24
anaphora 22, 27, 46, 83, 84, 93, 94, 97, 104-106, 167, 196, 201, 202
anaphora, direct 28, 49, 50, 54, 78, 81, 82, 92-95, 98, 104, 105, 107, 108, 109, 111, 114, 117, 124, 152, 154, 197, 200
anaphora, associative 46
anaphora, short-distance 101, 102
anaphora, long-distance 102
anaphoric 44-47, 49, 50, 51, 65, 73, 84, 92-95, 98-100, 103-106, 109, 152, 199, 200
anchor 108, 163, 197
antecedent 44, 146, 163, 164
appellative 80
attribute 22, 24, 26, 58, 59, 60, 61, 62, 63, 68, 69, 71, 75, 153, 162, 183, 208
attrition 38

B

bare noun 24-27, 29, 30, 32, 33, 54, 81, 88, 89, 93-98, 105, 109, 111, 117, 119-125, 130, 131, 134, 137, 138, 139, 142, 152, 172, 175, 176, 177, 179, 181, 183, 190, 202, 204
bleaching 38-40, 53
bondedness 37, 42

C

cardinality 31, 54
case 14, 22, 23, 58, 62-65, 67, 90, 139-144, 153, 198, 200, 203, 204
cline 13, 42, 43, 66, 73
clitic 12, 21, 41, 42, 43, 66, 67, 72, 73, 78
cliticization 34, 42, 57, 66, 67, 69, 72, 74, 77, 78, 204
cliticized 21, 23, 41, 58, 73, 75
cliticization gap 42
comparison 32, 53, 122, 123, 175-179, 184, 189
completeness 115, 143
concrete 172
conditional 32, 53, 175, 180-182, 184, 187
connotation 53, 199
constituent 42, 67
context, negative polarity 31, 32, 53, 110, 130, 160, 161, 174, 184, 189, 203
context, bridging 40, 43, 44, 49, 50, 53, 54, 171, 172
context, switch 40, 43, 44, 49, 50, 51, 53, 54, 171
contrast 24, 30, 32, 46, 51, 52, 60, 61, 64, 72, 74, 76, 109, 139, 155, 158, 159, 160, 190
conventionalization 43, 44, 49, 51, 53, 54
cointextual 146-151, 154, 198, 203
count 32
countable 33, 82, 86, 174, 177

dative 91, 92, 125, 138, 139, 140, 141,
declension weak adjectival 12, 59, 207
definiteness effect 96
deictic 27, 28, 35, 44, 45, 46, 49, 65, 67, 70
deixis 46, 49, 57, 64, 107, 132, 154, 196
demonstrative 11, 12, 13, 15, 21, 23, 28, 37, 38, 39, 40, 41, 44, 45, 46, 47, 48, 49, 50, 51, 54, 55, 56, 57, 58, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 82, 86, 92, 93, 94, 100, 101, 102, 104, 107, 109, 111, 115, 131, 143, 144, 154, 196, 197, 199, 200, 201, 202, 205
denotation 53, 200
discourse referent 27, 53, 161, 162, 166, 167, 168, 169, 170, 171, 174, 193, 197, 198, 200, 201, 202
distribution 11, 41, 42, 43, 71, 72, 76, 80, 90, 122, 131, 139, 145, 151, 152, 173, 176, 184
dividua 24, 32

e
emphasis 36, 102, 104, 150, 200,
enclitic 67, 75

f
fulpark, runic alphabet 14

G
gender 22, 32, 33, 41, 90, 199
generic 22, 28, 29, 30, 49, 51, 53, 54, 72, 78, 89, 121, 133, 134, 148, 150-152, 154, 196, 202, 203
genericity 53, 199
genitive 42, 111, 138, 139, 140, 141, 145, 146
gradience 55
grammaticalization chain 35, 37, 38, 44, 45, 52, 55, 56, 201

H
hypo-determinative language 96

I
inalienable possession 28, 87, 116, 118, 119, 120, 121, 151, 201
inclusiveness 29
individua 32
individualizing 59, 61
individualization 145
inference 43, 44, 51
innovation 58, 62, 74, 188, 194

L
lexicalized phrase 24, 25, 37, 105, 122, 129, 150, 152, 206
lifespan 53, 161, 170, 174, 193, 199, 200

M
mass 32, 33, 82, 86, 173, 177, 189
morphologization 58

N
negation 31, 36, 37, 38, 53, 65, 66, 130, 149, 161, 175, 179-180, 181, 184, 189
negative polarity 31, 32, 53, 54, 110, 130, 160, 161, 174, 184, 187, 189, 203
nominative 12, 62, 64, 90, 99, 100, 125, 139, 140, 141, 142, 144
non-agentive 97
INDEX 225

O
opacity 51, 160, 174
ostensive 45

P
partitive 191
persistence 43, 161, 170
persistent 54, 167, 168, 200
possessive 28, 39, 45, 46, 47, 86, 87, 107, 109, 110, 111, 115-121, 131, 145, 146, 196, 201
postposed 23, 37, 58, 67, 69, 71, 72, 75, 76, 78, 201
preposed 12, 23, 24, 69, 84, 100, 194

Q
quantification 53, 199
quantifier 33

R
reanalysis 68
reference 30, 32, 39, 50, 51, 53, 54, 71, 73, 80, 84, 92, 94, 99, 102, 104, 105, 132, 143, 161, 163, 171, 181, 182, 196
reference, unique 22, 49, 152, reference, generic 22, 28, 29, 134, 152, 154, referential distance 93

S
scope of negation 31, 53, 161, 179, 181
suffix 12, 67

T
topic 96, 158, 165-168, 171, 174
topical 96, 169, 201
trigger 43, 108, 112-114

U
uncountable 33, 202
unique 22, 23, 28, 48, 49, 51, 59, 61, 78, 80, 89, 90, 104, 107, 121-125, 127-132, 135, 136, 143, 144, 146, 148, 152, 154, 163, 190, 196, 204
uniqueness 49, 135, 136, 196
Gramatykalizacja (nie)określoności w języku szwedzkim

Streszczenie

Współczesny język szwedzki jest językiem rodzajnikowym: w swej historii wykształcił zarówno rodzajnik określony jak i nieokreślony. Grupy imienne określone używane są do opisania referentów wiadomych (Weinsberg 1983), znanych lub identyfikowalnych przez ich wcześniejszą obecność w tekście (anafora) lub poprzez odnoszenie się do zjawisk i rzeczy istniejących tylko w jednym egzemplarzu (unikatów). Grupy imienne nieokreślone odnoszą się do referentów po raz pierwszy używanych w tekście. Oprócz tego język szwedzki posługuje się grupami imiennymi bezrodzajnikowymi w zleksykalizowanych połączeniach wyrazowych oraz z rzeczownikami masowymi (dividua, z łac. ‘dający się dzielić’).

We współczesnym języku szwedzkim określoność w grupie imiennej (z przymiotnikiem) jest wyrażona trzykrotnie: przez rodzajnik prepozycyjny, słabą odmianę przymiotnika i rodzajnik postpozycyjny, np. den kort resa ‘(ta) krótka podróż’. Pod względem reprezentacji morfologicznej brak jest symetrii między grupą określoną a nieokreślona, w której pojawia się wyłącznie prepozycyjny rodzajnik nieokreślony, en kort resa ‘krótka podróż’.


Rozdziały 1-3 stanowią tło głównego badania. Przedstawiono w nich kategorię określoności we współczesnym języku szwedzkim, periodyzację języka szwedzkiego, materiał badawczy oraz zastosowane metody. Rozdział 2 przedstawia teoretyczne podstawy badania, procesy gramatykalizacji rodzajników określonych i nieokreślonych. Jak się wydaje, procesy te są zbieżne w niespokrewnionych językach w zakresie swych źródeł (rodzajnik określony jest gramatykalizacją elementu deiktycznego, np. zaimka wskazującego, rodzajnik nieokreślony gramatykalizacją liczebnika ‘jeden’). Modele gramatykalizacji rodzajników przedstawione zostały w Lyons 1975 i Greenberg 1978 (rodzajnik określony) oraz Givón 1981 i Heine 1997 (nieokreślony). Dla rodzajnika określone-
GRAMATYKALIZACJA (NIE)OKREŚLONOŚCI W JĘZYKU SZWEDZKIM

... go początkiem gramatykalizacji jest użycie zaimka deiktycznego w tekście dla oznaczenia anafory; dla nieokreślonego pierwszym etapem jest użycie liczebnika 'jeden' w funkcji prezentacji nowych referentów tekstowych. Ponadto rozwój rodzajnika nieokreślonego wydaje się być częściowo uwarunkowany rozwojem określonego – jest przesunięty w czasie wobec tego ostatniego. Z kolei rozwój rodzajnika nieokreślonego zawęża możliwe użycia określonego. Mamy zatem do czynienia z dwoma niezależnymi lecz wzajemnie się warunkującymi procesami.

Rozwój słabiej odmiany przymiotnika i klitycyzacja zaimka wskazującego mają miejsce w historii przedpiśmiennej, zatem w rozdziale 3 przedstawione zostały etymologie wszystkich wykładników określoności, hipotezy na temat przyczyn klitycyzacji oraz świadectwo ich rozwoju w epoce runicznej (800-1225), przed spisaniem dłuższych tekstów pismem łacińskim.

Rozdział 4 stanowi szczegółową analizę użycia -IN w tekstach staroszwedzkich z lat 1225-1526. Użycia te zostały podzielone na tekstowe oraz pozatekstowe. ‘Wiadomość’ (określoność) tych pierwszych opiera się na wcześniejszym użyciu w tekście lub na inferencji z innej informacji językowej. Są to zatem użycia anaforyczne, bezpośrednie lub pośrednie (te ostatnie są w anglojęzycznej literaturze przedmiotem określane też jako associative anaphora). Użycia pozatekstowe -IN obejmują unikaty (np. słońce) oraz użycie generyczne. Szczegółowa analiza wykazuje, że użycia tekstowe -IN pojawiają się już w najstarszych zabytkach piśmienniczych. W pierwszych tekstach nie są jeszcze obligatoryjne, a ich dystrybucję ograniczają taki czynnik jak np. odległość między wyrażeniem anaforycznym i jego antecedensem (im dłuższa tym większe prawdopodobieństwo, że -IN zostanie użyte) lub rodzaj relacji między nimi (przy nazwach części ciała bardziej prawdopodobne jest użycie zaimka dzierżawczego). W okresie 1375-1450 tekstowe użycie -IN staje się obligatoryjne, a użycie pozatekstowe staje się coraz bardziej powszechne. Okres 1450-1526 to czas stabilizacji -IN we wszystkich wymienionych funkcjach; użycie w kontekstach generycznych nie jest jeszcze w pełni regularne. Gramatykalizacja zaimka wskazującego w rodzajnik określony przebiega zatem od tekstowych do pozatekstowych funkcji, a jej pierwszym etapem jest przeniesienie wskazywania w przestrzeni na wskazywanie w tekście (anafora).

Rozdział 4 zawiera ponadto szczegółową analizę nietypowego z punktu widzenia współczesnej szwedczyzny użycia -IN, w tym przy rzeczownikach bonde ‘chłop’ w kodeksach prawnych oraz z abstrakcyjnymi rzeczownikami nazywającymi cechy np. fatikdom ‘bieda’ w prozie religijnej.

Rozdział 5 podobnie szczegółowej analizie został poddany liczebnik ‘jeden’ i jego dalszy rozwój w rodzajnik nieokreślony EN. Proces ten jest późniejszy niż omawiana wcześniej gramatykalizacja rodzajnika określonego; w okresie 1225-1375 EN pojawia się w funkcjach rodzajnikowych sporadycznie i tylko w prozie religijnej, w najstarszych tekstach, prawniczych, EN pełni wyłącznie funkcję liczebnika. Pierwsze nieliczebnikowe użycia EN ograniczają się do inicjalnych podmiotów wprowadzających nowego referenta, do którego w dalszym tekście pojawiają się kolejne odniesienia. Takie użycie EN możemy zatem nazwać kataforycznym. Podobnie jak w przypadku rodzajnika określonego, również gramatykalizacja rodzajnika nieokreślonego rozpoczyna się zatem od jego użyć tekstowych. W dalszych etapach rozwój obserwujemy rozszerzenie użycia EN na podmioty nie inicjalne oraz na inne części zdania; EN wprowadza nadal nowego referenta jednak niekoniecznie takiego, który będzie wspomniany w dalszej części tekstu. W ostatnim etapie rozwoju EN może być użyte w kontekstach tzw. negative polarity, po negacji, w zdaniach warunkowych, po czasownikach typu vilja ‘chcieć’ czy w porównaniach.

Rozdział 6 zawiera podsumowanie wyników. Wskazuje też na współzależności między obiem gramatykalizacjami. Gramatykalizację rodzajnika nieokreślonego rozpoczyna użycie EN dla prezentacji nowych referentów – to użycie pojawia się w momencie, w którym już obowiązkowo oznaczana jest anafora, czyli oznaczanie referentów już znanych. Oba mechanizmy wiążą się z tekstowym użyciem rodzajników. Wyniki wskazują też na to, że zainicjowana gramatykalizacja rodzajnika nieokreślonego ogranicza gramatykalizację rodzajnika określonego, który w braku przeciwwagi w postaci EN by mógł rozwinać się w rodzajnik specyficzny a nie tylko określony.
Gramatykalizacje obu rodzajników przypadają na okres intensywnych zmian w języku szwedzkim, towarzyszy im zanik fleksji nominalnej i kategorii przypadka, stabilizacja szyku i wykształcenie obowiązkowego podmiotu (szw. *subjektstvång*). Wszystkie te zmiany są wzajemnie zależne i wzajemnie się warunkują. Intensywnym zmianom językowym w tym okresie towarzyszą gwałtowe zmiany społeczno-kulturowe: intensywna chrystianizacja i kształtowanie się szwedzkiej państwowości.