Relative Clauses in Ælfric’s Catholic Homilies:

a quantitative study

Malcolm Golmann
Degree Project
Linguistics
HT 2007
Supervisor: Nils-Lennart Johannesson
ABSTRACT

The aim of this degree project has been to examine, analyze and describe which intra-linguistic factors influence how relative clauses are formed in Old English.

The key to successfully performing the task of identifying which factors influences the relative causes is to examine how these factors are distributed among the relative clauses in the text. The main focus of this investigation thus was to investigate how the grammatical features of the antecedents of the relative clauses in Old English were distributed. By analyzing a text sample of the work of the Old English writer Ælfric, taken from the Dictionary of Old English Corpus at the University of Toronto, also known as the Toronto Corpus, several features of the antecedent will ideally become evident as influencing factors.

The relative clauses that are found to be relevant for this investigation in the Ælfric text sample have been categorized and analyzed in order to identify any grammatical pattern that could indicate which factors influence how relative clauses in Old English are formed. The findings have been analyzed according to quantitative and statistical principles, and the chi-square test has been employed to verify the statistical significance of these findings. By doing this some linguistic factors have been verified as influencing factors.

**Key words:** Old English, antecedent, relative clause, boundary marking strategy, case marking strategy, combination strategy, the indeclinable particle þe, demonstrative pronoun, contingency table, statistically significant, Ælfric, Catholic homilies.
Table of Contents

1. Introduction 3

2. Primary Sources 4

3. Secondary Sources 5

4. Method 6
   4.1 Method Problems 7

5. Results 9
   5.1 Gender of antecedent 9
   5.2 Number of antecedent 13
   5.3 Case of antecedent 14
   5.4 Animacy of the antecedent 15
   5.5 Type of Noun Phrase as antecedent 16
   5.6 Syntactic function of relativised Noun Phrase 18
   5.7 Case of relativised Noun Phrase 19
   5.8 Distance 20
   5.9 Summary 21

6. Conclusion 22

7. References 24
1. Introduction

The purpose of this degree project has been to investigate relative clauses in Old English. There are three different ways in which a relative clause can be formed in Old English. This aspect of Old English has been somewhat neglected since the scholars that concern themselves with Old English usually describe the system that is involved in forming relative clauses, but not the linguistic features that govern how these relative clauses are formed. We simply do not know with any real certainty which linguistic factors in Old English influence how relative clauses are formed. The Old English textbooks and grammars that currently are available deal very briefly, if at all, with any intra-linguistic or syntactic factors that might influence how a relative clause was formed in Old English. The three different ways to form a relative clause in Old English are illustrated in example (1) below.

(1)  a. Ic eom se liflica hlaf: þe of heofenum asta h, \textit{ÆCHom I, 2} 192.82  
    I am the vital bread, which descended from heaven, (Thorpe, p 35)

b. Þa leasgewitan þa ledon heora hacelan ætforan fotum sumes georges cnihtes: Se wæs geciged saulus. \textit{ÆCHom I, 3} 200.52  
    The false witnesses then laid their coats before the feet of a young man who was called Saul. (Thorpe, p 49)

c. Eala þu betleem iudeisc land ne eart ðu nateshwon wacost burga on iudeiscum ealdrum: of þe cymþ se heretoga. se þe gewylt & gewissað israhela folc. \textit{ÆCHom I, 5} 217.17  
    Ah thou Bethlehem, Judæan land, thou art in no wise meanest of cities among the Jewish princes: of thee shall come the ruler who shall rule and govern the people of Israel. (Thorpe, p 79)

In example (1a) the indeclinable particle \textit{þe} introduces the relative clause, in example (1b) the demonstrative pronoun \textit{se} introduces the relative clause, and in example (1c) the demonstrative pronoun \textit{se} in combined with the indeclinable particle \textit{þe} to introduce the relative clause.

This degree project aims to perform a linguistic investigation that will examine which intra-linguistic factors in Old English influence the choice of how to form a relative clause in any of the three ways illustrated above. Hopefully this linguistic investigation will be able to answer the questions: Is any grammatical or syntactical feature in Old English an influencing factor for how the relative clause is formed? Is it possible to identify the influencing factors by examining the distribution of the relative clauses?
2. Primary Sources

To be able to perform a linguistic investigation of the relative clauses in Old English it is important to choose the primary material carefully. It soon became evident that the writings of the monk Ælfric are suitable for this task. Ælfric (ca 955-1010) was the abbot of the monastery of Cerne and one of the major writers who wrote in Old English. Ælfric’s production is both vast and varied, and as the Old English scholar and Ælfric expert Malcolm Godden points out, some of his writings were intended for a large and in some cases uneducated audience. Among the large production of religious texts, such as bible commentaries and Latin grammar books, Ælfric wrote a series of literary speeches that could either be used as sermons at mass in the monasteries or could be read as texts by any reader who had access to them and knew how to read. These sermons are called Ælfric’s Catholic Homilies and were written around 990 AD. The Catholic Homilies were very influential both in and after the Old English period and greatly circulated. The Catholic Homilies thus seem ideal for the task of examining relative clauses in Old English.

Ælfric’s Catholic Homilies are to be found in the Dictionary of Old English Corpus at the University of Toronto, which is also known as the Toronto Corpus of Old English. This is in effect a database that aims to have at least one copy of every Old English text read into it. This database has been purchased by the English Department at the University of Stockholm and is available on CD-ROM. The linguistic investigation of this degree project has used the Toronto Corpus entry for Ælfric’s Catholic Homilies as its primary material.

In some corpuses the text has been tagged, which means that each word has been labeled with its syntactic information and such. This is not the case of the Toronto Corpus. The text of Ælfric’s Catholic Homilies that has been examined is a plain text since this corpus is not tagged. All examples that are used in this degree project have been sought through text reading. This has been supported by the computer program Highlight that was used to mark words such as the particle þe and the demonstrative pronoun se. The problem with this program is just that it also marks all the cases of þe and se when they are used in other grammatical contexts. Hence, the Highlight program marks a lot of cases that are not relevant to this linguistic investigation. The quest for finding relevant examples of relative clauses in Ælfric’s Catholic Homilies has also been helped by using Benjamin Thorpe’s modern English translation from 1837. The size of the text sample that has been examined consist of the

---

1 For a full introduction to Ælfric and his world, see Godden, page xxii - xliv.
2 For a more extensive account of Ælfric’s Catholic Homilies, see Godden, page xxix - xxxvi.
3 The web address is as follows: http://www.doe.utoronto.ca/pub/webcorpus.html.
preface and the six first Homilies, roughly some 20 000 words /forty pages of Old English from the Corpus and about 100 pages from Thorpe’s translation. All the examples of Old English are taken from the Toronto Corpus. The examples are presented as they appear in the Corpus, with no editing or correction done. All examples of Old English are also provided with an identification tag in the form of the short title with page and line reference from the Toronto corpus, as well as a translation from Thorpe (1837) with a page reference.

3. Secondary Sources
The secondary material that this degree project refers to is mainly the work of Peter Baker and Nils-Lennart Johannesson. Baker (2007: 47-48) briefly describes how there are three main ways which relative clauses can be formed in Old English. He does not examine the linguistic factors that motivate the different ways to form relative clauses in Old English. All the same his work is very helpful when one needs to approach the basics of Old English grammar. Perhaps one should also briefly mention that Elizabeth Closs Traugott formulates a more detailed description of the usage of relative clauses in her chapter in *The Cambridge History of the English Language*. This description is more detailed than Baker’s, and in many ways very much more interesting, but is still just a description of the system, not an analysis of which factors motivate it.

Nils-Lennart Johannesson goes slightly further than his peers in Johannesson (2007:66-68). Therein he refers to the three ways of forming relative clauses in Old English as strategies. He also gives a detailed analysis of how the three different major strategies are used to form relative clauses in Old English. The first of these major strategies he calls the boundary marking strategy. He explains that the boundary marking strategy is used by marking the beginning of the relative clause with the indeclinable particle *þe* and that this strategy in some ways corresponds to how *that* is used in relative clauses in modern English, see example 1a above. The second major strategy for forming relative clauses in Old English he calls the case marking strategy. In this strategy the distal demonstrative pronoun *se* is used as a relative pronoun to introduce the relative clause, see example 1b. He then continues by showing that this strategy in many ways corresponds to how the relative *wh*-pronouns are used in Modern English. The third and final major strategy for forming relative clauses in Old English he calls the combination strategy. In this third strategy the methods of the two previous strategies are combined, where the distal demonstrative *se* in its complex form is combined with the indeclinable particle *þe* to introduce the relative clause, see example 1c.
This degree project will implement and discuss the categories described by Johannesson (2007), but also the concepts of linguistic investigation in Johannesson (1990). The terms and methodology concerning quantitative research that is described therein is perfectly suited for the type of intra-linguistic investigation that this degree project aims at performing. The method of this degree project is largely dependent on the concepts of analysis and evaluation that are described in Johannesson (1990).

4. Method

Quantitative research methods are often used in systematic and scientific investigation of grammatical features and their relationships. Since this degree project aims to investigate which linguistic factors influence how relative clauses are formed in Old English, a quantitative method seemed appropriate to this task. The implementation of this method has been divided into three steps. The first step consists in identifying the relevant instances where relative clauses appear in the selected Ælfric Homilies and to copy them into a separate computer file. The second step of the work consists in classifying all the findings that have been gathered into the computer file and sorting them into groups and subgroups according to their grammatical features. The third and final step is to analyze the findings based on a statistical evaluation tool known as the chi-square test.

To perform the identification and gathering process of relative clauses that is relevant to this linguistic investigation it is necessary to compare the cases that were found by the Highlight program with the corresponding examples in Thorpe’s translation from 1837. All findings of relative clauses that can be confirmed as relevant are then entered and numbered, both the Old English relative clause and the corresponding Modern English translation into a table. This proved to be both time consuming and a bit problematic since there were some problems both with the Highlight program and Thorpe’s translation.

Since a relative clause by default functions as a post-modifier to a noun phrase it seemed that the most effective way to classify the material was by identifying different features of the noun phrase. In Old English each noun phrase is marked for gender, number and case, so these three groups are a good starting point for the classification process. Also the relationship between the noun phrase and the relative clause should be the basis for some of the categories for classification. Johannesson describes how the Noun Phrase and the relative clause interact as follows:

[T]he relative clause is added to a Noun Phrase (the antecedent) with the same reference as a Noun Phrase within the relative clause (the relativized NP).
Structurally the relative clause is inserted as a postmodifier in the antecedent Noun phrase. (Johannessen 2007:66)

Essential in the above quote is that the noun phrase that the relative clause is referring to is labeled as the antecedent of the relative clause. It was thus decided that the categories of classification would describe features of the antecedent and of the relative clause. The first five categories specify the properties of the antecedent: the antecedent gender (male, female or neutral), the antecedent number (singular or plural), the antecedent case (nominative, accusative, dative or genitive), the animacy of the antecedent (animate or inanimate) and finally what type of noun phrase it is (pronoun, noun headed NP or zero headed NP). The following categories focus on the properties of the relative clause: the type of relativizer that is used (i.e. which of the three major strategies is employed), the syntactic function of the relativized noun phrase (subject, object or complement in preposition phrases), the case of the relative pronoun if it is used (only applicable to the case marking strategy and the combination strategy). The final category marks if there is any separation between the antecedent noun phrase and the relative clause (yes or no). The relative clauses that are found in the selection of Ælfric’s Catholic Homilies that are examined in this degree project will be categorized according to the above-mentioned categories.

4.1 Method Problems

Throughout the process of identifying and classifying the relative clauses a few problems arose. Certain relative clauses defied the standard classification and had to be interpreted in a slightly unconventional manner. Example (2) below illustrates one type of problem that occurred a few times.

(2) Đa gesceafta þe sind þwyrllice gépuhte: hi sind to wrace gesceapene yfeldædum. ÆCHom I, 6 230.177

The creatures that are thought monstrous have been created for punishment of evil deeds. (Thorpe, p 103)

The antecedent gesceafta which means creatures is sometimes considered as having neutral grammatical gender and sometimes as having feminine grammatical gender⁴. In this investigation the occurrences of gesceafta could be interpreted either way, but they have been recorded as feminine throughout the investigation for practical reasons, since creatures are

---

⁴ Clark Hall, and Meritt, (1960, page 251).
animate and the neuter gender often denotes objects and non-living things. Another problem could arise with spelling variation such as in example (3).

(3) Da hyrdas þæ wacodon ofer heora eowde on cristes acennednesse. 

ÆCHom I, 2 193.102

The shepherds that watched over their flock at Christ’s birth. (Thorpe, p 36)

In example (3) the relativizer þæ could indicate several possible interpretations, one would be that þæ simply is a spelling error, where þæt should be the correct version; another interpretation would be that þæ is a spelling variant of þe. Since the antecedent hyrdas is masculine, þæ is considered to be is a spelling variant of þe throughout the investigation. A more difficult problem in the investigation part of the degree project could be how to interpret cases such as example (4).

(4) & þes deofol þe is gehaten antecrist. þæt is gereht ðwyrlic crist. 

ÆCHom I, (Pref) 175.78

And the devil, who is called Antichrist, which is interpreted, Opposition-Christ. (Thorpe, p 4)

This example is problematic since the clause þæt is gereht ðwyrlic crist has the antecedent antecrist. The problem is that antecrist is masculine and þæt is the demonstrative pronoun for neutral antecedents, not masculine. The best way to approach this problem is to disregard Thorpe’s translation and to interpret the clause as a parenthetical main clause, where the translation would be: "that [word] is interpreted Opposition-Christ". In this way we do not have to interpret þæt as having a masculine antecedent. This means that example (4) is removed from the investigation and not counted as a relative clause.
5. Results
The first and most important task when we try to analyze our findings is to identify which factors influence which in Old English sentences. To do this it is necessary to establish what part of the material constitutes the independent variable(s) and what part of the material constitutes the dependent variable. The general idea is that the dependent variable is influenced, or governed, by the independent variables. Throughout this process of analyzing the gathered data the working hypothesis will be that the dependent variable is the strategy for forming relative clauses (or relativizer, depending on how you choose to present relativization); and that the independent variables are here any grammatical phenomena related to the text which may possibly affect the choice of strategy. In order to examine the validity of the above statement and to analyze how the antecedent influences its relative clause the gathered material have been divided into several subgroups, or categories, as previously described in the method discussion.

5.1 Gender of antecedent
The first category to be examined is the grammatical gender of the antecedents. This category has three natural subgroups; masculine, feminine, and neuter. But we must also deal with the plural pronouns in this section, which is a little complicated. This last group needs to be included since Old English did not distinguish gender for plural pronouns. So the Old English plural pronoun þa which means those can equally refer to something that has masculine, feminine or neutral grammatical gender. It is therefore necessary to include the plural pronouns as a group of their own since they are not marked for gender in any way. The masculine antecedents were found to be by far the most frequent in all three strategies. The feminine, neutral and plural pronoun are somewhat evenly distributed in relation to each other.

In order to distinguish if the grammatical gender of the antecedent plays an important role in determining how relative clauses are formed in Old English we need to examine its frequency. Johannesson defines frequency as a measurement of how often an item (in this case an antecedent with a specific grammatical gender) occurs in a text corpus (Johannesson, 1990:85.) We must also bear in mind that we try to establish the frequency of the different variants of the antecedent gender. This means that we must examine how the three strategies are distributed over the different grammatical genders. In order to perform this task it is necessary establish their relative frequency. Johannesson defines the concept of relative frequency as follows:
If we want to determine the frequency of occurrence of a particular variant, then we should relate the number of occurrences of the variant not to the number of words in the text, but to the number of occurrences of the variable in question. In other words, the answer to the question ‘how often’ the variant occurs must be based on the proportion of the actual occurrences of the variant to the total number of its potential occurrences in the text (Johannesson 1990:90).

In the case of gender distribution among the antecedents in Old English it would not really clarify anything to identify that Ælfric uses 90 masculine antecedents and 23 feminine ones when he forms a relative clause with the indeclinable þe as its relativizer, (that is, when Ælfric employs the first strategy). Instead these figures must be related to the total number of masculine and feminine antecedents that have been found in the examined material, just as Johannesson suggests in the above paragraph, and then we must gather them into a so called contingency table.

So the 90 masculine antecedents that are to be found in the first strategy must be related to the total number of occurrences of masculine antecedents, just as the feminine antecedents must be analyzed in relation to the total amount of feminine antecedent. By doing this we can hopefully identify some kind of pattern that will clarify how relative clauses are formed in Old English. The distribution of the grammatical gender of the antecedents is gathered into the contingency table below. In Table 1 below þe represents the first strategy, the boundary marking strategy, se represents the second strategy, the case marking strategy and se þe represents the third strategy, the combination strategy:

Table 1: Distribution of the strategies over the gender of the antecedents.

<table>
<thead>
<tr>
<th>Gender</th>
<th>þe</th>
<th>se</th>
<th>se þe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>90 (66 %)</td>
<td>22 (16 %)</td>
<td>24 (18 %)</td>
<td>136</td>
</tr>
<tr>
<td>Feminine</td>
<td>23 (73 %)</td>
<td>7 (21 %)</td>
<td>2 (6 %)</td>
<td>32</td>
</tr>
<tr>
<td>Neuter</td>
<td>32 (86 %)</td>
<td>5 (14 %)</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>Plural pronoun</td>
<td>30 (99 %)</td>
<td>1 (1 %)</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>Σ</td>
<td>175</td>
<td>35</td>
<td>26</td>
<td>236</td>
</tr>
</tbody>
</table>

After analyzing how the strategies are distributed over the genders a few things are discernable. First and probably foremost is the fact that there is almost no variation concerning the distribution of the plural (genderless) pronouns. The majority, as illustrated in
example (5), favor the first indeclinable strategy, except in one case, which is illustrated in example (6).

(5)  Þa ðe his leasungum gelyfað þam he arað. ÆCHom I, (Pref) 175.95
Those who believe in his leasings, those he honors. (Thorpe, p 7)

(6)  He gesceop gesceafta þa ða he wolde. ÆCHom I, 1179.16
He created those creatures that he would (Thorpe, p 11)

Please note in example (5) that ðe is a spelling variant of þe. Example (6) on the other hand is problematic because it deviates so sharply from how the other plural pronouns interact with their relativizer. The antecedent in example (6) is actually þa and the relativizer is the declined ða. In this example Thorpe's translation is no guide to how the interpretation since þa typically is followed by þe. Another thing that stands out in the above table is the low frequency of feminine antecedents in the third strategy, only two occurrences, and the complete absence of neuter antecedents in the same category. The feminine antecedents in the combination strategy function just like the masculine ones in the text. If we compare example (7), which has a masculine antecedent, christ, and (8), which has a feminine antecedent, yld, no clear difference in how the antecedent interacts with its relative clause can be discerned.

(7)  he getacnode crist se ðe is soð yrfrënuma þæs ecan fæder. ÆCHom I, 2 192.66
he betokened Christ who is the true heir of the eternal father (Thorpe, p 33)

(8)  Eadug is heora yld: seo þe ða gyte ne mihte crist andettan. ÆCHom I, 5 220.96
Happy is their age, which could not yet acknowledge Christ. (Thorpe, p 85)

No clear gender difference can be discerned by comparing the masculine antecedents with the feminine ones, so the key to understanding how grammatical gender influences the choice of strategy for forming relative clauses in Old English is probably found in how they are distributed. It is also hard to evaluate if there is any gender related difference between the two examples (7) and (8) since the total number of feminine cases in this category is very low. It might also very well be that there simply are too few feminine antecedents in Ælfric’s text to evaluate their influence on the relative clause.

When we apply the chi square test to how the strategies for forming relative clauses are distributed over the grammatical gender in the above contingency table we get the result that the distribution is too low in certain categories to successfully perform the chi square test. There are simply not enough cases of feminine and neuter antecedent to ascertain if the distribution is statistically significant. Also one notices that there is very little variation between how the genders of the antecedents are distributed in the first strategy. We have to
approach the grammatical gender of the antecedent in a different manner to be able to evaluate its relevance to relative clauses.

In order to evaluate how the strategies are distributed among the genders it becomes necessary to combine the marking strategy (example 1b) and the combination strategy (example 1c) into one group. This meant that findings such as example (9) which employs the marking strategy will be counted together with findings such as example (7) to make one single category.

(9) Þonne cymð se antecrist se bið mennisc mann & soð deofol ÀECHom I, (Pref) 175.73

Then anti-christ shall come, who is human man and true devil (Thorpe, p 5)

We now have a division between the first strategy that only employs the indeclinable ðe as its relativizer and the second and third strategies that in some way employs a declinable demonstrative pronoun to form their relativizers. The new contingency table looks as follows:

Table 2: Distribution of the strategies over the gender of the antecedents, where the strategy using the indeclinable particle ðe is one group, and the other two strategies that use a demonstrative pronoun is combined into one group.

<table>
<thead>
<tr>
<th>Gender</th>
<th>ðe</th>
<th>se / se ðe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>90</td>
<td>46</td>
<td>136</td>
</tr>
<tr>
<td>Feminine</td>
<td>23</td>
<td>9</td>
<td>32</td>
</tr>
<tr>
<td>Neuter</td>
<td>32</td>
<td>5</td>
<td>37</td>
</tr>
</tbody>
</table>

| Σ            | 144 | 60        | 204   |

In Table 2 above the plural pronouns have been excluded from the analysis since they formally are not marked for gender. The chi-square test can still be used for the remaining examples, but the observed differences in the distribution of the two strategy groups is not really statistically significant, but very close (chi-square=5.82, p=0.0545). If we analyze Table 2 it seems that the indeclinable first strategy is favored when the antecedent has a grammatical gender that is feminine or neuter, and more strongly in the case of a neuter antecedent. Examples (10) and (11) respectively illustrate cases with feminine and neutral antecedents.

(10) Þa wæs hwæðre an mægð þe næfre ne abeah to nanum deofulgylde. ÀECHom I, 1 186.222

There was, nevertheless, one family which had never bent to any idol, (Thorpe, p 25)
The case marking and combination strategies, which has been combined into a single group in table 2, can on the other hand be assumed to be favored by an antecedent that has a masculine grammatical gender. It is most probable that gender would be a significant influencing factor if a larger material had been investigated. Now we will repeat the above procedure of analysis for all the main categories, one by one, in order to establish if there are any that actually influence the choice of strategy for forming relative clauses.

5.2 Number of antecedent

The next category to analyze is the grammatical number of the antecedent. In the contingency table below we can see how the strategies for forming relative clauses are distributed over the grammatical number of the antecedent. The plural pronouns are henceforth always included in the figures.

Table 3: Distribution of the strategies over the grammatical number of the antecedents.

<table>
<thead>
<tr>
<th>Number</th>
<th>þe</th>
<th>se</th>
<th>se þe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td>104 (65 %)</td>
<td>31 (19 %)</td>
<td>25 (16 %)</td>
<td>160</td>
</tr>
<tr>
<td>Plural</td>
<td>72 (95 %)</td>
<td>3 (4 %)</td>
<td>1 (1 %)</td>
<td>76</td>
</tr>
<tr>
<td>Σ</td>
<td>176</td>
<td>34</td>
<td>26</td>
<td>236</td>
</tr>
</tbody>
</table>

The difference in the distribution of strategies over different values of the grammatical number of the antecedent is statistically significant (24.198, d.f = 2, p<.001). This leads to an actual analysis of how these figures are relevant to the choice of strategy for forming relative clauses in Old English.

When we analyze the distribution in the above table we can see a sharp contrast from 65 to 95 % in the use of the plural for the first strategy. This indicates that the first category, that has the indeclinable þe as its relativizer, is strongly favored when Ælfric wrote sentences that had plural antecedents, like example (11), since the overwhelming majority of these antecedents employ this strategy. In both the second and the third strategy the numbers are reversed, indicating a sharp contrast in the use of plural antecedents for these two strategies. In the second strategy the relative frequency of 19 % when the antecedent is singular is in sharp contrast to the 4 % when the antecedent is plural. In the third strategy the difference is even clearer since the percentage is contrasted from a relative frequency of 16 % when the
antecedent is in the singular, like in example (7) and (8), to an almost non-existent 1 % when the antecedent is plural, as in example (6). Also it should be noted that example (6) is problematic in other ways as well, as stated earlier. This could indicate that this case is some form of aberration and that the figure for number distribution concerning the combination strategy should be even lower. It also becomes evident that the second and third strategies, that both to some extent use a demonstrative pronoun to form relative clauses, are favored when the antecedent is singular.

The grammatical number of the antecedent can safely be said to be an influencing factor in the choice of strategy for forming relative clauses in Old English. The analysis of how the grammatical number of the antecedent is distributed among the three strategies suggests that the fist strategy was mainly used for plural antecedents and that the other two strategies were mainly employed when the grammatical gender of the antecedent is singular.

5.3 Case of antecedent
The next category is the antecedent case. A distinct feature of Old English is its use of the five cases, nominative, accusative, genitive, dative and instrumental to inflect nouns, pronouns and adjectives to signal their function in sentences and clauses. Peter Baker describes how the Old English case system is used, and explains how the three cases subjective, objective and possessive in modern English is a remnant of this system (Baker 2007:35ff). The figures below illustrate how the strategies for forming relative clauses are distributed over the grammatical case of the antecedent. This category seemed most promising during the gathering stages of the degree project. It seemed probable that the antecedent case would somehow influence the choice of relative strategy since case plays such a central role in Old English. This was in fact incorrect. Table 4 below shows that no clear pattern could be discerned in how the three strategies for forming relative clauses in Old English was distributed over the case of the antecedent:

Table 4: Distribution of the strategies over the case of the antecedents.

<table>
<thead>
<tr>
<th>Antecedent case</th>
<th>þe</th>
<th>se</th>
<th>se þe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>81 (72 %)</td>
<td>19 (17 %)</td>
<td>12 (11 %)</td>
<td>112</td>
</tr>
<tr>
<td>Accusative</td>
<td>33 (73 %)</td>
<td>8 (18 %)</td>
<td>4 (9 %)</td>
<td>45</td>
</tr>
<tr>
<td>Genitive</td>
<td>12 (71 %)</td>
<td>3 (18 %)</td>
<td>2 (12 %)</td>
<td>17</td>
</tr>
<tr>
<td>Dative</td>
<td>50 (81 %)</td>
<td>4 (6 %)</td>
<td>8 (13 %)</td>
<td>62</td>
</tr>
</tbody>
</table>
One problem with Table 4 is that it is not really possible to distinguish if the case of the antecedent is an influencing factor or not since we cannot perform the chi-square test. One thing that is interesting in Table 4 is that the indeclinable particle þe has a relative frequency of 81% when it has a dative antecedent, and 71/73% when its antecedent has other cases. The dative case is therefore a good candidate for further investigation. The best way to further investigate the dative case is to combine the three other cases into a larger non-dative group, and thereby compare distribution of the strategies for forming relative clauses in the two new groups. By doing this we receive a distribution that the chi-square test accepts, which is illustrated in Table 5 below:

### Table 5: Distribution of the strategies over the case of the antecedents, where the antecedents with non-dative cases is combined into one group, and the antecedents with dative case is one group.

<table>
<thead>
<tr>
<th>Antecedent case</th>
<th>þe</th>
<th>se</th>
<th>se þe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-dative</td>
<td>126 (72 %)</td>
<td>30 (17 %)</td>
<td>18 (10 %)</td>
<td>174</td>
</tr>
<tr>
<td>Dative</td>
<td>50 (81 %)</td>
<td>4 (6 %)</td>
<td>8 (13 %)</td>
<td>62</td>
</tr>
</tbody>
</table>

The difference in the distribution of strategies over different values of the antecedent case is still not statistically significant (4.381, d.f = 2, p > .100). Now that we have been able to perform the chi-square test we know that the difference is not significant. The conclusion has to be that the antecedent case does not influence how relative clauses are formed in Old English.

### 5.4 Animacy of the antecedent

The next category involves the animate or inanimate status of the antecedent, i.e. whether the antecedent of the relative clause refers to a living being or not. The distribution how the strategies for forming relative clauses are distributed over animate versus inanimate antecedents in the contingency table below seems rather even but a few peculiarities should be noted, mainly that the combination strategy seems to be favored when the antecedent is animate.
Table 6: Distribution of the strategies over the animacy of the antecedents.

<table>
<thead>
<tr>
<th>Animacy</th>
<th>ḫe</th>
<th>se</th>
<th>se ḫe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animate</td>
<td>114 (74 %)</td>
<td>20 (13 %)</td>
<td>21 (13 %)</td>
<td>155</td>
</tr>
<tr>
<td>Inanimate</td>
<td>62 (77%)</td>
<td>14 (17 %)</td>
<td>5 (6 %)</td>
<td>81</td>
</tr>
<tr>
<td>Σ</td>
<td>176</td>
<td>34</td>
<td>26</td>
<td>236</td>
</tr>
</tbody>
</table>

The difference in the distribution of strategies over different values of the animacy of the antecedent was shown to be non-significant (3.399, d.f = 2, p>.100). All the same the numbers in the above table indicates, or hints, that the animacy of the antecedent is important and should perhaps be examined in future and larger linguistic investigations that deal with relative clauses in Old English.

5.5 Type of Noun Phrase as antecedent

The next category for analysis concerns the type of noun phrase as antecedent. A relative clause is typically a post-modifier to the head of a noun phrase, as in example (12), but in some cases the noun that functions as the head of the phrase is replaced by a pronoun, as in example (13), or even nothing in some cases, as in example (14). These last cases are called zero-headed noun phrases.

(12) ofsleað ealle þas iudeiscan ealdras þe ic on cweartene beclysde.  
ÆCHom I, 5 222.154
kill all the jewish elders whom I have confined in prison (Thorpe, p 87)

(13) þæt se is healicost se ḫe ðone martirdom æfter gode ast ealde.  
ÆCHom I, 3 201.81
that he is the most exalted who suffered martyrdom next to god (Thorpe, p 51)

(14) & se ilca se ḫe ableow on adames lichaman & him forgeaf saule. se ylca forgyfð cildum saule & lyf on hyra moder innoðe. þonne hy gesceapene beoð.  
ÆCHom I, 1184.173
And the same who blew into Adam’s body, and gave him a soul, that same giveth a soul and a life to children in their mother’s womb, when they are created. (Thorpe, p 21)

In Old English the majority of the noun phrases are headed by nouns. It is more uncommon that noun phrases are headed by pronouns in Old English, and zero-headed noun phrases seem
outright rare. The contingency table below shows how the strategies for forming relative clauses are distributed over noun phrase heads in the examined material.

Table 7: Distribution of strategies over the types of noun phrase of the antecedents.

<table>
<thead>
<tr>
<th>Type of NP</th>
<th>þe (70%)</th>
<th>se (17%)</th>
<th>se þe (13%)</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun-headed NP</td>
<td>117</td>
<td>28</td>
<td>20</td>
<td>165</td>
</tr>
<tr>
<td>Pronoun</td>
<td>54 (86%)</td>
<td>4 (6%)</td>
<td>5 (8%)</td>
<td>63</td>
</tr>
<tr>
<td>Zero-headed NP</td>
<td>5 (62%)</td>
<td>2 (25%)</td>
<td>1 (13%)</td>
<td>8</td>
</tr>
</tbody>
</table>

Σ 176 34 26 236

In this category the findings are problematic. The chi-square test rejects the findings due to the low number of zero-headed noun phrases and pronoun headed noun phrases is certain columns. One way to proceed further in the analysis of how the type of noun phrase influences how relative clauses were formed is to merge some of our smaller groups into larger ones. This can be done by contrasting the pronouns with the non-pronouns. Since the pronouns is the group that stand out from the other two groups, it seems logical to form one group from the pronouns and to combine the other two groups, the noun-headed noun phrases and the zero-headed noun phrases, into one single group. If we do this we receive the following contingency table:

Table 8: Distribution of the strategies over the type of noun phrase of the antecedents, where antecedents with non-pronoun heads are combined into one group, and antecedents with pronouns as heads is one group.

<table>
<thead>
<tr>
<th>Type of NP</th>
<th>þe (70%)</th>
<th>se (17%)</th>
<th>se þe (13%)</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-pronoun</td>
<td>122</td>
<td>30</td>
<td>21</td>
<td>173</td>
</tr>
<tr>
<td>Pronoun</td>
<td>54 (86%)</td>
<td>4 (6%)</td>
<td>5 (8%)</td>
<td>63</td>
</tr>
</tbody>
</table>

Σ 176 34 26 236

The difference in the distribution of strategies over different values of the noun phrase type is now statistically significant (6.043, d.f = 2, p<.050). This statistic figure suggests that pronoun headed antecedents favor the indeclinable particle þe. Reversely, antecedents that are headed by non-pronouns seem to favor the strategies that form their relativizers with demonstrative pronouns, according to the percentages. The next step in examining the noun
phrase type’s relationship to the three strategies is to look into the distribution of the genders over the two categories of pronoun and non-pronoun; perhaps the pronoun/non-pronoun contrast is due to an over-representation of neuters in the pronoun group (or the other way around). The plural pronouns are not included in table 9 since they are not marked for gender.

The new table looks as follows.

Table 9: Distribution of the strategies over the type of noun phrase of the antecedents, where the groups of the previous table are also divided after the gender of the antecedent.

<table>
<thead>
<tr>
<th>Type of NP</th>
<th>Gender</th>
<th>þe</th>
<th>se /se þe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-pronoun</td>
<td>M</td>
<td>69 (63 %)</td>
<td>41 (37 %)</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>23 (72 %)</td>
<td>9 (28 %)</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>30 (97 %)</td>
<td>1 (3 %)</td>
<td>31</td>
</tr>
<tr>
<td>Pronoun</td>
<td>M</td>
<td>21 (81 %)</td>
<td>5 (19 %)</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>2 (33 %)</td>
<td>4 (67 %)</td>
<td>6</td>
</tr>
<tr>
<td>Σ</td>
<td></td>
<td>145</td>
<td>60</td>
<td>205</td>
</tr>
</tbody>
</table>

The difference in the distribution of strategies over different values of the non-pronoun antecedents is statistically significant (13.520, d.f = 2, p<.005). The chi-square test cannot be used for the pronoun antecedents. By examining Table 9 we can see that the boundary marking strategy seems favored when the antecedent is a neuter non-pronoun, and that the case marking strategy and the combination strategy seems to be favored when the antecedent is a masculine non-pronoun.

5.6 Syntactic function of relativised Noun Phrase

The next category for examination is the syntactic function of the relativised NP within the relative clause. Throughout the examined material of this degree project the syntactic function of the relativised noun phrase within the relative clauses has been found to consist of a rather large distribution of subjects and objects, a few complements to preposition phrases and one determiner. The distribution of the strategies for forming relative clauses over the syntactic function is displayed in the below contingency table.
Table 10: Distribution of the strategies over the syntactic function of the antecedents.

<table>
<thead>
<tr>
<th>Syntactic function</th>
<th>þe</th>
<th>se</th>
<th>se þe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>128</td>
<td>25</td>
<td>22</td>
<td>175</td>
</tr>
<tr>
<td>Object</td>
<td>39</td>
<td>6</td>
<td>4</td>
<td>49</td>
</tr>
<tr>
<td>PP comp</td>
<td>9</td>
<td>2</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Determiner</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

| Σ                  | 176  | 34   | 26    | 236  |

Just as in the antecedent case the syntactic function turned out to be very problematic. For some of the functions the numbers were too low to be useful for evaluation and even if one divided the groups into subjects and others the chi-square test showed an error probability of 25%. With such high error probability and such low distribution in certain subgroup such as preposition phrase complements and the determiner, we are forced to reject the syntactic function of the relativised noun phrase within the relative clause as an influencing factor.

5.7 Case of relativised Noun Phrase

The next category involves the case of the relativised noun phrase. An important feature of Old English is that the relativizer agrees with the number and gender of the antecedent. This obviously does not apply to the first strategy, the boundary marking strategy since it uses the indeclinable þe as its relativizer. This means that the case of the relativised noun phrase imposes itself on the relativizer in the second strategy, the marking strategy, and the third strategy, the combination strategy. For instance, if the relativized element is in the accusative case the demonstrative pronoun is normally declined in the accusative, as in example (15).

(15) Ne mæg ic minne feond lufian. þone þe ic dæghwamliche wælreowne togeanes me geseo. ÆCHom I, 3 204.167
    I cannot love my foe, whom I see daily bloodthirsty against me (Thorpe, p 55)

One hope was that the use of the nominative and accusative case for the relativised noun phrase would show a pattern when the case marking strategy would be employed versus the combination strategy. One hypothesis was that the case marking strategy would be favored when the relativised noun phrase was marked for the accusative and that the combination strategy would be favored when the relativised noun phrase was marked for the nominative. This notion could not be ascertained by the findings of the examined text, which are illustrated in Table 11 below:
Table 11: Distribution of the strategies over the case of the relativised noun phrase of the antecedents.

<table>
<thead>
<tr>
<th>Case of relativised NP</th>
<th>se</th>
<th>se þe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>25 (53 %)</td>
<td>22 (47 %)</td>
<td>47</td>
</tr>
<tr>
<td>Accusative</td>
<td>10 (77%)</td>
<td>3 (23 %)</td>
<td>13</td>
</tr>
</tbody>
</table>

The difference in the distribution of strategies over different values of the case of the relativised noun phrase was shown to be non-significant (2.360, d.f = 1, p>.100). This of course means that we cannot accept these figures as reliable for this linguistic investigation. All the same a few comments are warranted concerning the relative frequency of the distribution in the above table. When the case of the relative clause in the marking strategy is in the accusative case the 77 % percentage is in contrast to the considerably lower 52 % when the case is in the nominative. Consequently the percentage is reversed in the combination strategy, where the 23 % percentage is contrasted to the accusative 47 % in the nominative. This distribution could indicate that the accusative case favors the marking strategy for the relative clause and that the nominative favors the combination strategy. This cannot be certified since the error probability is too high to be acceptable but the relationship between the case of the relative clause and the choice of strategy should be investigated examined in future and larger linguistic investigations that deal with relative clauses in Old English.

### 5.8 Distance

The final category for analysis examines if there is any separation between the antecedent, the head of the noun phrase, and its post-modifier, the relative clause. In most cases the noun phrase remains intact but in a few cases the head and the post-modifier are separated by an intervening element of some kind. This is illustrated in the example (16) below:

(16) þa wearð he to deofle awend. se ðe ær wæs mare engel geworht. *ÆCHom I, 1 180.57]*

Then became he changed to the devil, who before was created a great angel.

(Thorpe, p 13)

In the above example the head of the noun phrase is *he* is and the relativizer is *se ðe*. Notice that the above example employs the combination strategy and that the antecedent and the corresponding relative clause are separated by *to deofle awend*. How the strategies for
forming relative clauses are distributed over separated and proximate post-modifiers is shown below:

Table 12: Distribution of the strategies over the distance of the antecedents to their post-modifiers.

<table>
<thead>
<tr>
<th>Separation</th>
<th>þe</th>
<th>se</th>
<th>se þe</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximate</td>
<td>146 (79%)</td>
<td>25 (13%)</td>
<td>13 (14%)</td>
<td>184</td>
</tr>
<tr>
<td>Distance</td>
<td>30 (58%)</td>
<td>9 (17%)</td>
<td>13 (25%)</td>
<td>52</td>
</tr>
<tr>
<td>Σ</td>
<td>176</td>
<td>34</td>
<td>26</td>
<td>236</td>
</tr>
</tbody>
</table>

The difference in the distribution of strategies over different values of the separation of the antecedent from its relative clause is statistically significant (14.776, d.f = 2, p<.001). In the first strategy that uses the indeclinable particle þe, the boundary marking strategy, we note that the percentage is very high when there is no separation between the antecedent and the relative clause, almost eighty percent, but in cases where the relative clause is separated from its antecedent the percentage is only around sixty percent. It thus seems that it is more common to use the indeclinable particle þe as relativizer when the relative clause is proximate to its antecedent. In the other two strategies, which both use a demonstrative pronoun to form their relative clauses, the situation seems less clear. In both strategies the percentages are higher when the relative clause is separated from its antecedent, specially the combination strategy. This suggests that it is more common to use a demonstrative pronoun as a relativizer (sometimes in combination together with the indeclinable particle þe) when the relative clause is separated from its antecedent.

5.9 Summary

Out of all the above categories, only four can be accepted as statistically significant: antecedent gender, number, noun phrase type and separation. Two other categories, the animacy of the antecedent and the case of the relative clause, are bordering on being statistically significant but had to be discarded all the same due to the fact that their error probability was too high. To further analyze the gathered material we could measure how distribution is influenced by its surroundings by performing the VARBRUL analysis. Johannesson has successfully used this analysis when he performed a similar linguistic investigation based on examination of contingency tables. He describes the use of the VARBRUL Analysis as follows:

\[5\text{Nearly so; cf. the discussion in section 5.1.}\]
What is needed for an analysis of the kind of data presented in the previous section is an instrument that allows us to measure the impact of each environment (or influencing factor) independently of all the other environments. Such an instrument is available in the form of the computer program VARBRUL, which predicts, on the basis of the observed data the individual contribution of each influencing factor to the probability that a variable rule (in this case, topicalization) will be applied. (Johannesson, 1989:118)

It is for practical reasons not possible to perform a VARBRUL analysis on the material that has been analyzed in the degree project due to the fact that it would be too time and space consuming to do so. But if one would be interested in expanding the analysis of the gathered data the VARBRUL analysis is clearly the way to proceed.

6. Conclusion

The aim of this degree project has been to perform a linguistic investigation that will examine which intra-linguistic factors influenced how relative clauses were formed in Old English. In this limited study, four grammatical features stand out as clear influence on how relative clauses were formed in Old English. Three of these are grammatical features of the antecedent, namely the gender, number and noun phrase type of the antecedent. The fourth factor concerns the question of whether the relative clause is separated from its antecedent, that is, if the post-modifier in the actual noun phrase is separated from its head. Thus, the boundary marking strategy can be said to be favored by an antecedent that is either a plural pronoun or a neuter pronoun, and by a relative clause that comes immediately after the antecedent. By contrast, the case marking and combination strategies can be assumed to be favored by an antecedent that is masculine, singular, and a full noun phrase, and by a relative clause that is separated from its head.

Another question that this degree project hoped to shed light on was which of the three ways to form a relative clause is the most common. By examining how the three strategies for forming relative clauses are distributed among the text samples that have been examined in this degree project, the answer must be that the most common way by far to form relative clauses in Old English is by using the indeclinable particle þe as relativizer, as the boundary marking strategy does. A methodical aspect of this project has also been to try to establish if it is possible to identify the influencing factors by examining the distribution of the relative clauses. The answer to this inquiry is that it is possible to identify influencing intra-linguistic
factors by analyzing their distribution in the examined text sample by using statistical evaluation methods. The conclusion is that it is established, at least to some extent, that the antecedent influences the choice of strategy for forming Old English relative clauses.
7. References


