

Grammaticalization of Complementizers in Old English Glosses

by

Catherine Mackowski

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved November 2010 by the
Graduate Supervisory Committee:

Elly van Gelderen, Chair
Karen Adams
Robert Bjork

ARIZONA STATE UNIVERSITY

December 2010

ABSTRACT

I investigate how complementizers, which connect subordinate clauses to the main sentence, develop from other parts of speech, namely prepositions and adverbs. This occurs by the process of grammaticalization, in which a word loses lexicality and gains grammatical function instead. I use computer-based corpus analysis to determine how often certain words are used as each part of speech in my selected texts, and whether they are accompanied by other grammatical words. I use two Old English glosses of the Latin gospels, the Rushworth and Lindisfarne glosses, in order to analyze possible diachronic and geographical differences between the texts. I demonstrate that prepositions develop into adverbs and thence into complementizers with the assistance of certain grammatical accessory words which are later lost. This occurs by the process of reanalysis, in which the language user interprets a word or phrase differently than before.

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
LIST OF ABBREVIATIONS.....	vi
CHAPTER	
1 INTRODUCTION AND OVERVIEW	1
1.0 Introduction	1
1.1 Choosing the Topic	2
1.2 The Texts.....	3
1.3 The Importance of Word Order	4
1.4 Grammaticalization and Reanalysis.....	5
1.5 The Origin of Complementizers	7
1.6 Proposal of Research.....	8
1.7 Summary	9
2 OLD ENGLISH AND ITS GRAMMAR.....	10
2.0 Introduction	10
2.1 History of Old English and its Dialects	10
2.2 The Dialects and History of My Texts.....	11
2.3 Old English Word Order	12
2.4 Ambiguities in Old English	14
2.5 Complementizers in Old English	15
2.6 Summary	17

CHAPTER	Page
3 SYNTACTICAL THEORY AND GRAMMATICALIZATION	18
3.0 Introduction.....	18
3.1 Universal Grammar and X-bar Theory	18
3.2 Movement.....	20
3.3 Grammaticalization.....	22
3.4 Grammaticalization of Complementizers.....	23
3.5 Dedicated Old English Complementizers	24
3.6 Reanalysis	24
3.7 Model: Preposition to Complementizer	26
3.8 Model: Adverb to Complementizer	29
3.9 A Problem with the Model.....	30
3.10 Summary	31
4 QUANTITATIVE ANALYSIS OF THE TEXTS	33
4.0 Introduction	33
4.1 Notes on the Data.....	33
4.2 Æfter	35
4.3 Ær	36
4.4 Butan	37
4.5 Mid	38
4.6 Oð.....	39
4.7 Wið	40

CHAPTER	Page
4.8 Summary	41
5 THE MODEL APPLIED TO THE DATA	42
5.0 Introduction	42
5.1 Non-initial Prepositional Phrase	42
5.2 Initial Prepositional Phrase	43
5.3 Reduced Prepositional Phrase	44
5.4 Complex Complementizer	45
5.5 Simple Complementizer	47
5.6 Adverbs	48
5.7 Summary	49
6 ANALYSIS AND DISCUSSION.....	51
6.0 Introduction	51
6.1 Æfter	51
6.2 Ær	52
6.3 Butan	52
6.4 Mid	52
6.5 Oð.....	53
6.6 Wið.....	53
6.7 Differences Between Texts	54
6.8 Trends and Conclusions	55
REFERENCES	59

APPENDIX	Page
A TOTAL OCCURENCES BY WORD	61
B THE RAW DATA	66

LIST OF TABLES

Table		Page
1.	Æfter and its Accessory Words	35
2.	Ær and its Accessory Words	36
3.	Butan and its Accessory Words	37
4.	Mið and its Accessory Words	38
5.	Oþ and its Accessory Words	39
6.	Wið and its Accessory Words	40

LIST OF ABBREVIATIONS

Acc(-word)	Accessory (Word)
Av	Adverb
AvP	Adverb Phrase
C	Complementizer (I generally use this to refer to the position in an X-bar structure, rather than the word that fills the position)
CP	Complementizer Phrase
Cz	Complementizer
D	Determiner
DP	Determiner Phrase
IP	Inflection Phrase (Layer)
J	Gospel of John
L	Gospel of Luke
MdE	Middle English
Mk	Gospel of Mark
MnE	Modern English
Mt	Gospel of Matthew
NP	Noun Phrase
OE	Old English
ON	Old Norse
P	Preposition

PoS	Part of Speech
PP	Prepositional Phrase
P, P-word	Accessory word beginning with the letter p
UG	Universal Grammar
V	Verb
VP	Verb Phrase
V2	Verb Second (word order)

CHAPTER 1

INTRODUCTION AND OVERVIEW

1.0 Introduction

In this introduction chapter I give a summary of my research problem. I explain my interest in the topic, give some background information on my texts and languages, the general theory of syntax, explain my methods of investigation, and state my research questions. Chapter Two gives a more thorough background in Old English (OE) language and grammar. I give a brief history of the language and dialects as they relate to my texts, and summarize the possible word orders of OE clauses and how they relate to determining the part of speech of the words I am investigating. Chapter Three gives a more thorough background into the syntactical aspects of my investigation. I give a brief explanation of X-bar theory and Universal Grammar (UG). I then discuss the theory of grammaticalization and how it changes a word from one part of speech (PoS) to another. I then present my step-by-step theory of how a preposition (P) grammaticalizes to an adverb (Av) and/or a complementizer (Cz). Chapter Four presents my raw data. I determine how many times each word appears as a given PoS, in what form (alone or in a phrase with one or more accessory words), and assess the relative frequencies of each. In Chapter Five I analyze my data. I apply my step-by-step model to the words and see whether the data supports each step. In Chapter Six I discuss my results. I assess where each word is on a grammaticalization cline of P → Av → Cz, and determine whether any particular text or dialect shows evidence

of more grammaticalization than another. The data itself is presented in the Appendices.

1.1 Choosing the Topic

I find Old English word order endlessly fascinating (and frustrating), with flexibility and exceptions to every rule. I was originally interested in landing sites for floating adverb phrases, which can be placed nearly anywhere in the sentence. This was complicated when I realized that an adverb is not always an adverb: just as in Modern English (MnE), the categories of adverb, preposition, and complementizer frequently overlap. How does a word in one category move into another? It is not precisely grammaticalization: while there is an increase in the abstract grammaticality of a complementizer as compared to an adverb or preposition, there is no diminution of phoneticity. In fact, some words have remained in multiple categories for hundreds of years. Is this a very slow transition from one to another, or is this a stable system? What quality or feature of the word or category of words allows it to be analyzed in multiple ways? Was there something pronounced or lexical in the Old English system, when the change was presumably just beginning, to signal the difference in function?

I decided to see if there was any diachronic change in the usage of certain words as certain parts of speech. This would require a relatively large text which could be dated and located with some degree of precision. In addition, it must be a prose text, due to the stylistic constraints of OE poetry that affect word order and the poetic tendency to have long parallel phrases. I chose to look at the Rushworth

and Lindisfarne Gospels due to the fact that they are surprisingly specific in their dates and locations. Essentially they make up two different dialectal versions of the same text, allowing easy comparison. I cannot think of anything else existing in two “versions,” except perhaps the West Saxon gospels.

1.2 The Texts

The OE texts I am looking at are the Rushworth and Lindisfarne glosses. They are both translations of the Latin texts of the gospels. The Lindisfarne glosses were made in the Northumbrian dialect in the late 10th century (Skeat, John ix). The Rushworth glosses were written about a hundred years later by two different scribes. The first scribe, Farman, glossed Matthew and three verses of John in his native Mercian dialect, and did so loosely rather than as a close copy. He started on Mark, and having gained access to the Lindisfarne glosses at this time, began copying them directly rather than making his own looser glosses as he had been doing before. Soon he passed the task to the second scribe, Owun, who continued copying the Lindisfarne glosses in Northumbrian through the rest of the gospels. Thus the text known as Ru1 (Farman’s section) is in Mercian, and Ru2 (Owun’s section) is in Northumbrian (Skeat, John xii).

Changes between the Rushworth and Lindisfarne glosses should indicate a difference that the scribe felt was significant enough to correct: the language had changed enough either between dialects or in the intervening years that the old text was no longer felt to be acceptable and needed to be corrected. There is likely to be little to no difference between Ru2 and Lindisfarne, the former being a

relatively close copy of the latter. However, the differences between Ru1 and Lindisfarne are likely to be much greater, due to the combination of difference in dialect, the passing of a century, and the looser translation of Farman compared to Owun. While I do not think the geographical or dialectal differences are likely to be great, the change of a hundred years should show a larger effect: late OE was rapidly evolving as it began to change into Early Middle English, and showing influence from French and Norse (Mitchell and Robinson 132-4).

1.3 The Importance of Word Order

It can be argued that glosses, as opposed to translations or original compositions, are not good indicators of OE word order and syntax, since they are interlinear translations, following the word order of the Latin text. Latin has a basically subject-object-verb word order, which is also the underlying word order of OE (before processes such as verb movement) (Haspelmath et al. 331). A Latin verb-final sentence is likely to result in a reasonably intelligible OE sentence when translated directly, without change in word order. Latin Czs, like OE Czs, occur at the head (left side) of their clause; a subordinate clause may occur at the beginning, middle, or end of the sentence, but in this text is found most often at the beginning. So the word order of a Latin sentence is natural to an OE reader/speaker, certainly more so than the word order of most OE poetry.

In any case, word order is not my main indicator of whether or not a word is a complementizer. A subordinate clause headed by a Cz must have its own verb, in addition to the verb of the main clause. A preposition would be followed

by a noun phrase (NP or DP) rather than a verb phrase; the sentence would only have one verb. Likewise an adverb phrase would not be followed by its own verb. A sentence with a Cz will have a rough order of Cz VP VP, or VP Cz VP; a sentence with an AvP or PP will only have one VP in it. Since I am only trying to differentiate Cz from Av from P, this is sufficient for my purposes. Latin, like OE, keeps its clauses together, i.e., one can easily tell where one clause ends and the other begins, since the words in one clause are not mixed with the words of another except in certain poetic devices of classical poetry which are not found in the Bible.

1.4 Grammaticalization and Reanalysis

Grammaticalization is a theory within the UG framework that can be used to describe how words change function and category from one part of speech to another. It is the process by which a word becomes less lexical and more grammatical or functional. As the word loses lexical and semantic content, it often also loses some phonological content, becoming shortened or unstressed. As it gains grammatical function it becomes more abstract (Hopper and Traugott 2-3). For instance, the auxiliary verb “have” was originally only a lexical verb, having the possessive meaning and not indicating tense or aspect: “I have a pie.” It later gained perfective and causative meanings: “I have a pie made.” These uses were then reanalyzed as an auxiliary verb, obligatorily accompanied by a finite verb, indicating tense or aspect: “I have made a pie.” The auxiliary verb is unstressed, and this may be further reduced to the clitic form “I’ve made a pie.”

Reanalysis is the process of reinterpretation by a speaker/listener of a language, whereby a word or phrase is understood, semantically or syntactically, differently by the hearer than was intended by the speaker (Hopper and Traugott 50). It is distinct from analogy, in which a new meaning is created based on resemblance to another; the original may be borrowed from another language. Language contact can lead to borrowing of grammatical usage patterns or to other, unrelated changes in a language (Heine and Kuteva 4-5). Also, not all reanalysis is grammatical: it may lead to lexicalization or changes in word order; grammaticalization is a subset of reanalysis (Hopper and Traugott 58-59). While reanalysis itself is not directional, grammaticalization is an accumulation of reanalyses in a specific direction (Roberts 48). In my model, I refer to reanalysis as the final step of the process of grammaticalization; it is the point, set up by earlier steps, in which the words actually change in meaning in the hearer's mind from a less to a more grammatical phrase.

Reanalysis is important because it is the smallest step of language change: an utterance with a certain meaning to the speaker has a slightly different meaning to the hearer (Hopper and Traugott 39). Reanalysis is internal to the listener and occurs in tiny increments, but an accumulation of reanalyses by many speakers will eventually result in significant language differences. As the speaker learns the language, he or she hears the language spoken by others (L1), interprets it according to the rules provided by his internal UG, and produces his or her own version of that language (L2) (Hopper and Traugott 40-41). L2 and L1 are of

course nearly identical, but there are subtle differences: a sentence that seems ungrammatical to a speaker of L1 may be acceptable to a speaker of L2, or a word that has a certain set of functions in L1 may have a different set of functions in L2. (L1 → L2 → L3 ... L99 → L100, so that L1 and L100 have noticeable differences.) Essentially, high lexicality + low grammaticality → low lexicality + high grammaticality.

1.5 The Origin of Complementizers

Most Czs develop from a different part of speech, usually adverbs (Av), prepositions (P), or sometimes determiners (D) (Hopper and Traugott 187-96). In Old English, most words that can function as both C and Av or P have the same form for both uses (i.e. they are not declined, shortened, or otherwise changed). This is similar to MnEng, in which the phrases “before dinner” and “before we ate dinner” are equally acceptable, and the word “after” has the same form as a P as it does as a Cz. However, when the word has a more grammatical function in OE (i.e. it is an Av or Cz), it is often accompanied by “grammatical words” such as *þa*, *þe*, or *þæt* (Mitchell and Robinson 88-89). Most of these grammatical words serve as complementizers or relative pronouns in their own right, and may have transferred this function to the former preposition.

Grammaticalization of Av and P to Cz consists mostly of reanalysis. For P developing into Cz, a word that previously (as a P) went before a DP or NP to set it off as a sub-unit of the sentence, now has the same function for a VP or IP. For Av, a word that previously indicated temporal/spatial location for a verb now

indicates the temporal/spatial relationship of the clause to the main sentence (Hopper and Traugott 85). This is a relatively shallow level of grammaticalization, so it seems right that just as there is only a little increase in grammaticality, there is no loss of phonological and little of semantic content. In fact, the addition of phonological content through grammatical words as grammaticality increases seems to go against the usual pattern of grammaticalization, although this additional content is reduced by the time of MdEng (Mosse 116-17) and completely gone by MnEng.

1.6 Proposal of Research

For this project, I have chosen several words from OE that can serve as prepositions, adverbs, and complementizers, with or without accompanying grammatical words. Using corpus analysis software, I count the occurrences of each word as each part of speech in my texts, and note whether they occur with grammatical words. I will attempt to answer several questions:

- Can a given word be used as multiple parts of speech without grammatical words?
- Are particular grammatical words used to turn one PoS into another? (i.e. P to Av, Av to Cz, P to Cz)
- Do the two texts show different patterns of PoS/grammatical words, indicating grammaticalization?
- Are these differences due to diachronic or geographical differences, or authorial choice?

- What do the patterns of PoS/grammatical words show about the process of reanalysis that changes one PoS into another?
- What are some possible paths/processes of reanalysis?

I believe that sentence-initial PPs were reanalyzed as AvPs and CzS. Some of their increased grammaticality may have been gained by their adjacency to grammatical words already having the role of Cz or Av. Reanalysis of an adjacent PP and Cz as a complex Cz would result in the old P gaining a new Cz function.

1.7 Summary

In the OE language, like MnE, certain words can function as multiple parts of speech. This seems to be an example of grammaticalization, in which words lose phonetic detail in exchange for increased grammatical function, but in this case they retain their old functions as well. I propose to investigate how this change occurs by measuring how often certain words are used as prepositions, adverbs, and complementizers. I measure the differences in usage between two texts that differ in date and dialect, but are otherwise very similar. I believe the change may be assisted by the use of other grammatical words which may co-occur with my selected words of interest.

CHAPTER 2

OLD ENGLISH AND ITS GRAMMAR

2.0 Introduction

In this chapter I will give a brief background of the Old English language as it relates to my research. I discuss the history of the language and its dialects, and how my texts are positioned in relation to these. I then discuss OE word order and sentence structure. Finally I cover how to determine which part of speech a given word is, and how this relates to word order.

2.1 History of Old English and its Dialects

OE is the language spoken in England from approximately 450 AD up through the Norman Conquest in 1066, after which it developed into Middle English and thence into the Modern English spoken today. OE itself developed from the West-Germanic languages brought by the successive invasions of Angles, Saxons, and Jutes in the 6th century, who established various warring kingdoms (Mitchell and Robinson 118-19). OE is generally divided into four dialects: Northumbrian, spoken north of the Humber river and into lower Scotland; Mercian, spoken in the midlands; West Saxon, spoken south of the Thames; and Kentish, spoken in Kent. West Saxon became the most politically important of the dialects after Alfred the Great unified the kingdoms in the late 9th century, and most extant OE writings are in West Saxon (Campbell 8-9).

Northumbrian and Mercian were probably the most innovative of the OE dialects because of the contact with the Norse-speaking Viking invaders in the

area (Mitchell and Robinson xi). The new Norse settlers, in order to communicate with their neighbors, learned a simplified version of OE. In turn, their neighbors picked up new words and even a few grammatical features (such as the pronoun system) from Old Norse (ON) (Mitchell and Robinson 133). OE's complex system of case endings and verbal endings was thus changed and simplified. Of course, these and many other changes occurred over a span of several centuries, and to a greater or lesser degree in the other dialects as well (Toon 60). However, it occurred most quickly in those areas with the greatest contact with another language, so that by the beginning of the Middle English period (about 1100), the dialects are readily distinguishable, especially the northern from the southern. (Mosse 2-3)

2.2 The Dialects and History of My Texts

The texts I am concerned with are written in Northumbrian and Mercian. Aldred, the glossator of the Lindisfarne gospels, and Owun, who copied most of them a hundred years later, both spoke Northumbrian. Farman, who copied the rest, spoke Mercian, although probably a northern variety which did not greatly differ from Northumbrian. (Harewood, the monastery at which the Rushworth gospels were made, is right at the border of the two dialects, and probably drew monks who spoke both or some mixture of the two.) (Skeat, John xiv)

There is also a time difference of about 100 years separating the Lindisfarne and Rushworth glosses (Skeat, Mark xi-xii). This combination of time and dialect differences between the texts is, as far as I can tell, a unique situation.

The Rushworth glosses are essentially copies of Lindisfarne, with one gloss expected to show differences of time and the other differences of time and dialect compared to the original. This arrangement of a control text and variations is ideal for the sort of quantitative study I perform here. However, this is a very small amount of time in terms of language change, and a very small difference in location, and may not be enough to produce noticeable changes. Noticeable differences might also be due to the scribes' personal taste or style, or the style expected of them by their superiors and the religious nature of their work.

2.3 Old English Word Order

Old English, like many other Germanic languages, has underlying SOV word order. This underlying order can go through many movements to produce sentences with several different surface level word orders (Kiparsky 152-53). Verb movement is mandatory in some situations, prohibited in some situations, and optional in others. Pronouns can move in some situations according to stress-based rules, and word order in poetry may be made more flexible to follow rules of alliteration that are not found in prose (Kiparsky 146). The three main types of sentences are verb-first, verb-second, and verb-final.

Verb-final is probably the most common order; it is mandatory in coordinate and subordinate clauses, and can sometimes be found in main clauses as well. It may be found in main clauses if no movement occurs at all; the CP layer is present but not filled (Kiparsky 142).

1. *Her Wulfred ærcebiscep pallium onfeng.*

Here Wulfred archbishop cloak.of.office received.

“In this year, archbishop Wulfred took office.” (Anglo-Saxon Chron. 804)

It is found in subordinate clauses when movement is blocked by a complementizer (obligatory to subordinate clauses) (Traugott 108) .

2. *ða ða ic to rice feng*

then when I to rule took

“then when I became king” (Alfred 5)

It is found in coordinate clauses when movement is blocked.

3. *Ac ic ða sona eft me selfum andwyrde ond cwæð*

But I then soon after my self answered and said

“But I soon answered myself and said” (Alfred 6)

Verb-second (V2) is found when movement does occur. With the verb moving to the Cz position near the beginning of the sentence, it is preceded only by the word/phrase in specCP, making the verb second (Kiparsky 142).

4. *þy ilcan geare lædde Ecgbryht cyning fierd on NorþWalas*

in.the same year led Ecgbryht king campaign on North-Welshmen

The same year, king Ecgbryht led a campaign against the North Welsh.

(Anglo-Saxon Chron. 828)

Verb-first sentences also occur. They are found when the verb moves to the Cz position, but specCP is empty, leaving the verb first (Kiparsky 142).

5. *Song he ærest be middangeardes gesceape*

Sang he first about middle-earth's shaping

First he sang about the creation of the world. ("Bede's Account of
Cædmon," Mitchell and Robinson 222)

2.4 Ambiguities in Old English

OE word order can sometimes be hard to determine. The positions of adverbs, adjective phrases, prepositional phrases, and pronouns can be quite flexible. This is due mainly to the highly synthetic nature of the language, which allows words far apart to be linked by their case endings, and does not depend on a specific word order to make sense of the sentence. A sentence that is technically verb-final may have its verb followed by so many other words and phrases (direct and indirect objects, prepositional and adverbial phrases, etc.) that the verb is actually closer to the beginning of the sentence than the end (Fischer et al 144). Phrases may also be moved to the front of the sentence by focusing or topicalization, especially noun and adverb phrases.

OE, like MnE, has some words that can be used as several different parts of speech without changing their form. (I exclude such homographs as lead-lead.) These are generally adverbs, prepositions, and complementizers (Mitchell and Robinson 83-88); examples include *until* (P and Cz), *where* (Av and Cz), *above* (Av and P), and *before* (all three). (Most OE texts refer to complementizers as "conjunctions," a term that encompasses both coordinating and subordinating conjunctions; the latter are generally identical to complementizers. I use the word

conjunction only to refer to coordinating conjunctions.) Other words can only be used as one part of speech, such as *if* (Cz), *of* (P), and *never* (Av).

For this project, I set out to choose words that I knew to function as all three parts of speech; my words also could not be wh-type interrogative words (were, when) or þ-type locative words (there, thence), which behave very differently. Wh-words are always fronted in interrogative sentences; this part of the grammaticalization process is not done by the choice of the speaker, as is the case with the topicalization of other adverbial or prepositional phrases. Words that were not originally prepositions (such as wh-words and locative þ-words) grammaticalize differently from those that were; they are not followed by pronouns and demonstratives that can be mistaken for grammatical words (see section 3.7). Of the many possible paths by which a word may grammaticalize into a Cz, I decided to investigate only the path of P → Av → Cz, in order to see the greater differences between the starting and ending points of the cline. The six words of interest that I investigate can all occur as P, Av, or Cz; they are: *æfter* (after, again, according to), *ær* (before), *butan* (but, except, without, unless, outside), *mid* (with, while, when), *oð* (up to, until), and *wið* (against, until).

2.5 Complementizers in OE

Word order is an important clue to the structure of the OE sentence; by locating the verb, one can usually tell whether or not it is a main clause. However, word order alone cannot distinguish coordinate from subordinate clauses, which can be a problem due to the OE tendency to string many sentences or clauses

together with a series of “and”s (Fischer et al. 53). Word order also has little significance when the text is a gloss, as is the case here. Although Latin and OE have generally similar word orders, since they are both highly synthetic languages with underlying SOV order, beyond these general tendencies they can often vary greatly. A sentence that feels well-formed to a native Latin speaker would probably be understandable but not very well-formed to an OE speaker, and vice versa.

However, both of these languages are head-left. This means that the head of any grammatical phrase is usually on the leftmost side (Santorini and Kroch ch. 5). (E.g. “the,” the head of a determiner phrase, is the leftmost word of “the cat,” and “in,” the head of a prepositional phrase, is the leftmost word of “in the hat.”) Since the complementizer is the head of the CP, which is nearly always the highest level in the sentence or clause, it is nearly always the first word in the sentence or clause (Santorini and Kroch ch. 5). (Exceptions include conjunctions, which appear to the left of the Cz, and occasional noun phrases, usually in the form, “Bob, when he came home, ate dinner.”) This is equally true in Latin, OE, and MnE.

In my research, therefore, my main method of determining part of speech is by the word’s position in the sentence relative to the verbs. A sentence with a complementizer will have two clauses and therefore two verbs. If the word in question appears in a sentence with only one verb, it cannot be a complementizer. A complementizer will never be at the end of a sentence; an adverb may be at the

end, and a preposition is rarely so (the exception is usually the translation of a Latin clitic, such as “mecum” producing “mec mið,” which is literally “me with”). A preposition will be followed almost immediately by a noun or determiner; other parts of speech rarely so. In the Latin text, a complementizer is usually followed immediately by a verb, especially clauses beginning with “cum.”

So, although the order of the OE sentence is useful in determining the structure of the sentence and the types of clauses involved, it is not absolutely necessary to determine whether a word is a P, Av, or Cz. This is mainly determined by the number of finite verbs in the sentence and the word’s position relative to them.

2.6 Summary

Old English had several dialects and was spoken for several hundred years, leading to various internal changes in the grammar of the language, which may proceed differently in the different dialects. In all the dialects, word order is relatively flexible. An OE sentence may be verb-initial, verb-second, or verb-final, the last being the most common. Even though my texts are glosses of a Latin original, the word order is still similar to the natural word order of OE, since both are head-left languages with underlying SOV order. This means that prepositions and complementizers will be at the beginning of their respective phrases, so that the only information needed to determine PoS for a given word is whether it is followed by a noun (preposition) or verb phrase, and whether there is only one VP in the sentence (adverb), or multiple (complementizer).

CHAPTER 3

SYNTACTICAL THEORY AND GRAMMATICALIZATION

3.0 Introduction

In this chapter I will give a brief summary of the grammatical theory necessary to my research. I give some background on the basic theoretical model of Universal Grammar and how it is represented and diagrammed. I then discuss grammaticalization, the process by which a word becomes less lexical and more grammatical, and how this relates to the development of prepositions into adverbs and complementizers. Finally I present my model of how such a development might take place in a language.

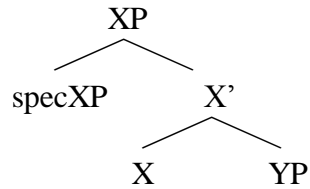
3.1 Universal Grammar and X-bar Theory

Generative grammar is the set of rules by which one can construct correct sentences in a language (Santorini and Kroch ch. 1). A sentence is divided into several layers, each containing a certain type of information about the sentence, which can be further divided. The lowest, the VP, contains the verb and its arguments and modifiers. The middle, the IP, contains information on tense, mood, and aspect. The highest, the CP, contains information on topic and focus, sentence or clause type (interrogative, emphatic, etc), and connection to other clauses (van Gelderen 251-252).

Sentences can be represented and diagrammed as tree structures using X-bar notation. Each unit or phrase has a head, a specifier, and a middle level where

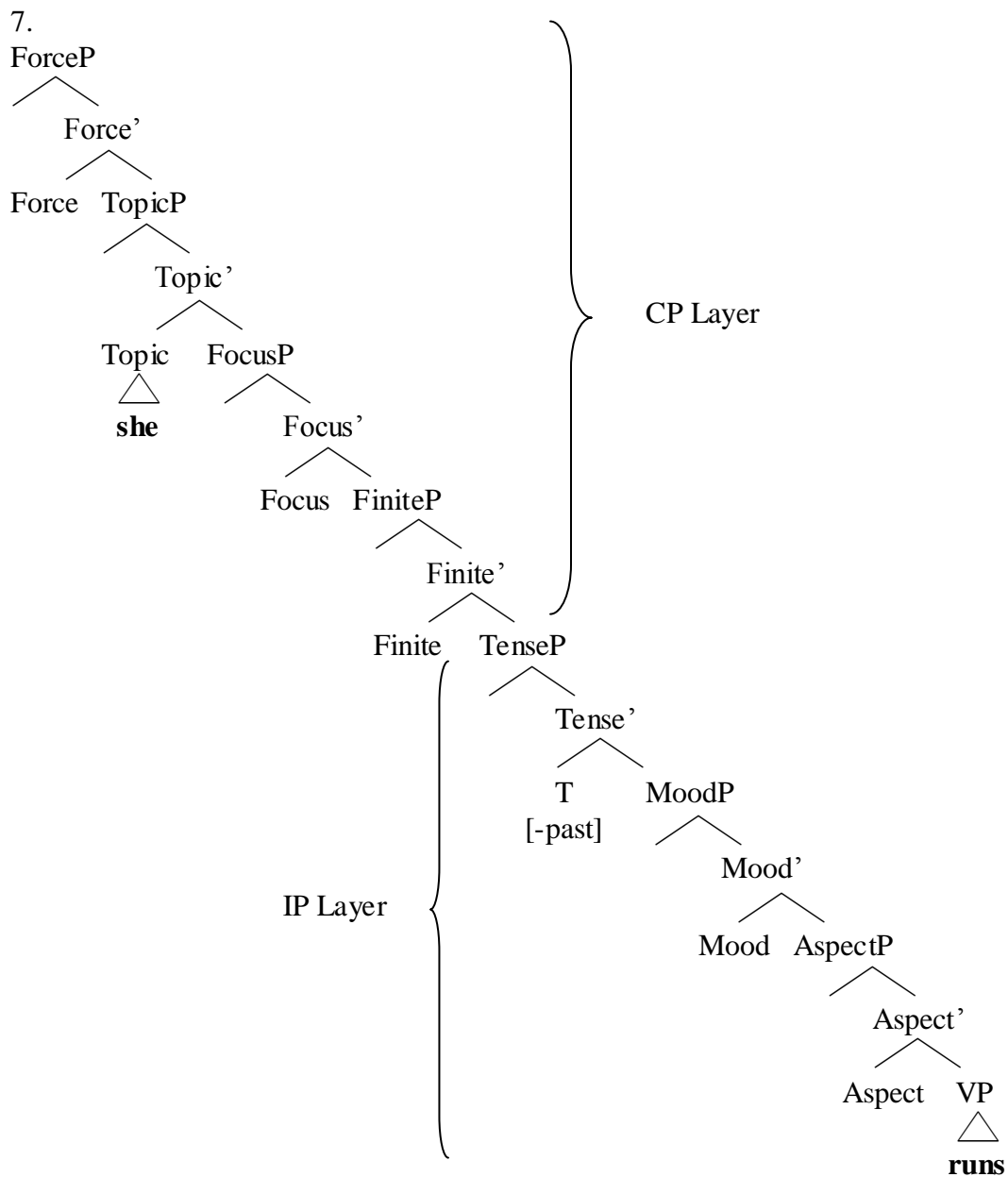
one can put arguments and modifiers (Santorini and Kroch ch. 4). This can be represented in tree form:

6.



Items can move between layers, mainly by head-to-head movement. For instance, the verb moves from V to I in order to create tense (or I moves down to V) and thence to C to represent sentence type. In another movement type, the subject of the verb moves from specVP to specIP in order to create agreement (Santorini and Kroch ch. 6).

Not all positions may be filled lexically. For instance, present tense is not usually expressed in MnE by a phoneme other than the verb itself, and so is represented in a diagram as [-past] or [+past]. I assume that all layers are present even when they are not filled by lexical items in a particular sentence, and can be seen as being filled by null items. Some layers are optional or moveable (such as Neg), and which layers are used can depend on which particular theorist's model is being followed. Thus a very simple sentence, such as "She runs," can be diagrammed with many layers, mostly filled with null items or nonlexical items.



However, I will show in my diagrams only those layers relevant to the discussion, namely the CP and VP layers (van Gelderen 136-7).

3.2 Movement

Obviously an item cannot move if it is blocked by the presence of something else in its landing site (or certain places in between). The most common element to be moved is the verb. The main restriction on movement,

thereby determining word order in a clause, is the fact that V cannot move to C if there is already a Cz at C. A Cz is mandatory in subordinate clauses in both OE and MnE, meaning that movement cannot occur and these clauses will be verb-final (Kiparsky 142-43). In MnE, this occurs only in specific situations. V-to-C movement is mandatory in questions, when the verb is brought forward to the beginning of the sentence (in the CP layer) (Santorini and Kroch ch. 11).

In OE, movement is generally required in the same places as MnE, and is optional in other main clauses (Kiparsky 142). V2 sentences are caused by V-to-C movement: the first item in the sentence is in specCP, and the next item, the verb, is in the C position, above and therefore to the left of the other elements of the sentence. Another example of movement is wh-raising, in which a question element (generally beginning with wh-) is brought to the front of the sentence to indicate that the sentence is interrogative (Santorini and Kroch ch. 11). Movement of other parts of speech can occur, the most relevant of which are topicalization and focusing (Kiparsky 144). These movements bring a word or phrase to the beginning of the sentence to emphasize its importance, and can act upon nouns, pronouns, adverbs, prepositions, and their phrases. Topicalization and focusing are notable in that they are often optional, consciously chosen by the speaker to give emphasis, and may not be required for grammaticality the same way V2 order or wh-raising might be.

3.3 Grammaticalization

Grammaticalization is the process by which lexical words or phrases, in certain situations, gain grammatical functions, and how grammatical words and phrases gain additional grammatical functions (Hopper and Traugott 1). Nouns, pronouns, verbs, adjectives, and most adverbs are lexical words: they are relatively concrete, and describe things, ideas and actions in the real world. On the other hand, prepositions, conjunctions, determiners, complementizers, and some adverbs are grammatical, or functional, words: they give more information about the relationships between the words themselves than about the real world. Lexical words tend to be longer and have more phonetic detail, and are usually created by invention or borrowing by speakers. Grammatical words tend to be shorter and less phonetic (e.g. they are often unstressed) and are created by development of older lexical words rather than by new invention (and borrowing of grammatical words is very rare) (Hopper and Traugott 4).

Grammaticalization tends to proceed along a cline from more lexical/less grammatical to more grammatical/less lexical. This cline is in most cases unidirectional: words seldom de-grammaticalize (Hopper and Traugott 6, 16). This semantic bleaching tends to occur at later stages of grammaticalization, outside the scope of this project, which focuses on the earlier stages; at these earlier stages the meaning tends to be shifted rather than deleted. For instance, a fully lexical verb (an independent, declinable word with relatively concrete meaning) may become an auxiliary verb (independent but indeclinable, and less

concrete) and thence a clitic (neither independent nor declinable, and relatively abstract). In most cases, there is some phonetic loss along with the lexical loss: a word tends to become more shortened or simplified as it becomes more grammatical (Hopper and Traugott 154-55). However, this does not always occur; forms can remain stable for a very long time, and phonological changes tend to happen very late in the process of grammaticalization, again outside the scope of this project (Hopper and Traugott 172).

3.4 Grammaticalization of Complementizers

Hopper and Traugott present a clause-combining cline (176-84), by which clauses become more closely connected. At one end is parataxis, in which two clauses are adjacent and presumably related, and are connected semantically. The two clauses are still independent syntactically and semantically. In the middle of the cline is hypotaxis, in which two clauses are interconnected but neither is completely dependent on the other. Clauses may be joined by some sort of conjunction or by verb chaining. Coordinate clauses, appositional clauses, and adverbial clauses are included in this type. The other end of the cline is subordination, in which the subordinated clause is completely dependent on its matrix clause, often acting as an argument of the matrix clause (Hopper and Traugott 183). Here I am concerned with the development of an adverbial clause from paratactic to hypotactic. Despite being joined by “subordinate conjunctions,” the adverbial clauses I discuss here never proceed as far as the stage of subordination.

Complementizers themselves may develop from nearly any part of speech: nouns, verbs, pronouns, adverbs, prepositions, affixes, and combinations thereof (Hopper and Traugott 184). They are often accompanied or assisted by deictics, which have a pre-existing connective function. In the case of an adverbial complementizer (the adverb clause is descriptive, rather than conditional, concessive, etc.), the deictic contributes connectivity, and the adverb contributes the manner of the connection. Only together can they express the relationship between the clauses that is intended by the speaker (Hopper and Traugott 185).

3.5 Dedicated Old English Complementizers

OE has two uninflected subordinators, *þe* and *þæt*, generally used for relative clauses. These may be used on their own or in combination with another word. In the case of *þe*, when placed with a preposition, this combination may be used as an adverb, and the part of speech of an ambiguous P or Av may be determined by the presence or absence of *þe* (Mitchell and Robinson 89-90). OE *þa* is also a subordinator meaning “when” (it can also appear as a clause-initial adverb meaning “then”). It generally connotes time or sequence of clauses (Mitchell and Robinson 86). For convenience I will be referring to these (along other small grammatical words) as *þ*-words, especially when speaking of broad cases in which any of these may be used to the same effect.

3.6 Reanalysis

Reanalysis is when the syntax of a sentence is interpreted differently than it was before, without actual modification to the sentence (Whitman 220). This

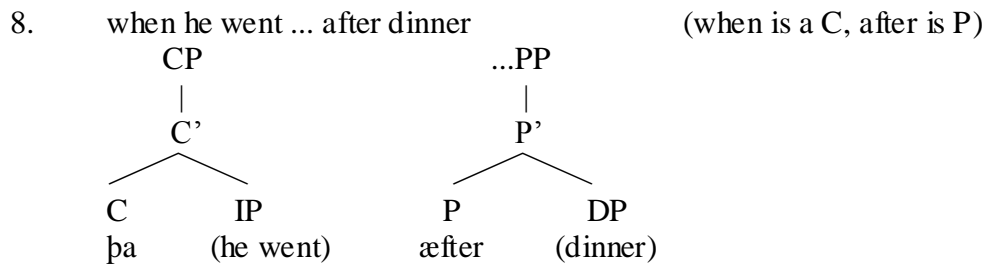
may be when the hearer understands a sentence differently from how it was intended by the speaker. It may also be when the same person is able to interpret a sentence in two different ways. The simplest of these is relabeling, in which words change in category (i.e. part of speech) without changing the structure of the sentence (Whitman 221). A change in label can trigger larger changes: a preposition may be followed by an NP; if it is relabeled as a Cz, it is now allowed to be followed by an IP/VP. Thus the relabeling of one item can lead to a more thorough reanalysis of the whole sentence.

However, this cannot be done all at once. Like most aspects of language change, it is done so slowly as to be unconscious (Pintzuk et al. 12). A speaker may be aware that some words are flexible in their PoS, but a speaker will not decide one day that a word which he previously treated as a preposition will now be treated as an adverb. Similar words might be confused or considered flexible, but only in the right situations. A flexible preposition will never be mistaken for a complementizer if it is at the end of a sentence, followed only by a noun phrase. I propose a model by which, moving in small steps of movement, mistakes, flexibility, and reanalysis, a preposition might eventually develop into an adverb or complementizer. My model has multiple paths: a preposition can develop into an adverb and either stop there, or continue on to become a complementizer. An adverb (not previously a preposition) may enter midway through the process and become a complementizer. However, like most processes of grammaticalization,

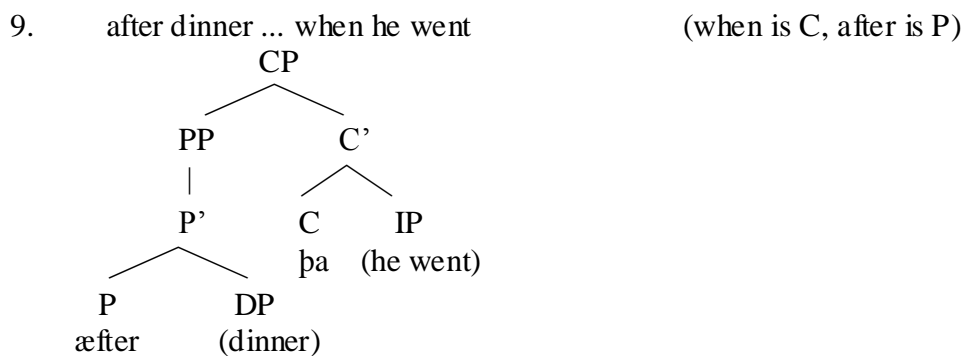
my model does not work in reverse: a complementizer will not become an adverb or preposition.

3.7 Model: Preposition to Complementizer

We begin with a sentence that has a pre-existing complementizer. The word of interest, “after,” is currently a preposition and happens to be located at the end of the sentence.

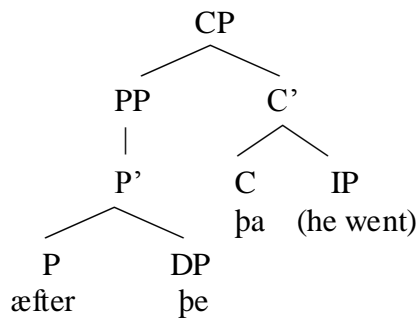


In order to possibly be mistaken for a complementizer, which only occurs at the beginning of the sentence, it is necessary to topicalize the entire prepositional phrase. Since “after” is the head of the PP and therefore at its beginning, when topicalized it will now be at the beginning of the sentence.



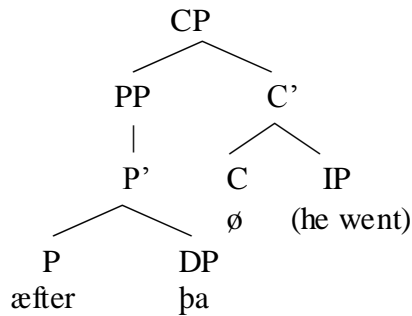
It would be possible for a listener to mistake one type of verb-final phrase for another: “a sentence is verb-final, and the phrase at the beginning is not a noun/pronoun, so it must be a complementizer of some sort, and so must be a subordinate clause.” However, it would be rather difficult to mistake an entire prepositional phrase for a complementizer, which is only one word. It would be easier if the PP was reduced, such as to the preposition plus a pronoun or adverb such as this/that or then/when. (such as β -words).

10. after which ... when he went (when is C, after is P)



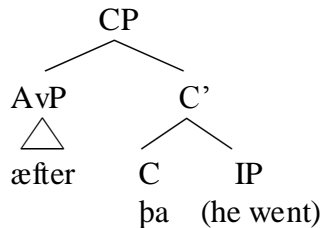
The β -word can then be re-analyzed as either adverbial β_a or relative β_e (or possibly even instrumental β_y). If it is analyzed as a pronoun β_a , it would produce an adverbial phrase “æfter þa” in the specCP position (since adverbial β_a always goes to specCP). The empty Cz position in the new C layer would trigger VtoC movement, resulting in a V2 word order, which is not the desired goal. However, it does leave us with an adverbial phrase in the specCP position.

11. after this (=these) he went (after is P, no C)



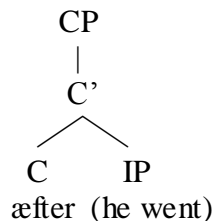
If the þ-word is analyzed as a complementizer þa, we get better results. þa is already a complementizer, the resulting phrase “æfter þa” would be analyzed in the Cz position and block VtoC movement.

12. after when he went (when is C, after is ambiguous)



At this point it is very easy to delete the small, unstressed þ-words, leaving only the more lexical (heavier) former-pronoun as the sole Cz.

13. after he went (after is C)

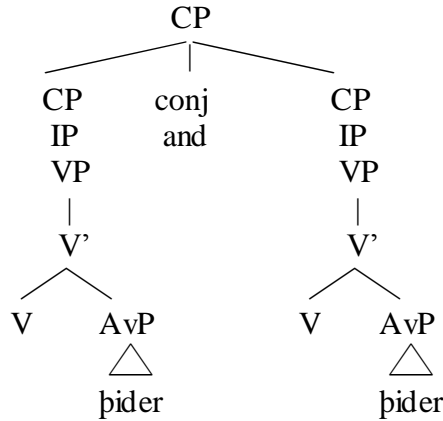


This final step, although arrived at in small increments, produces a sentence with a meaning entirely different from the original. The original meaning of “after X, then Y” is changed to “after Y,” a change so drastic that it is nearly impossible to do by any sort of accidental reanalysis.

3.8 Model: Adverb to Complementizer

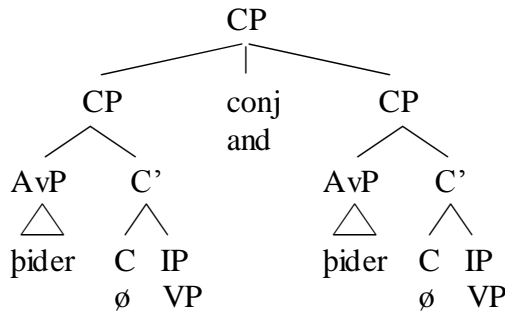
For this example, we begin with a word that is originally an adverb, never a preposition. Although this is not the case of the words I look at, it applies to many other OE words. As above, we begin with the Av occurring near the end of the sentence, but in this example we do not have a pre-existing Cz.

14. I go þider and you go þider (þider is Av)



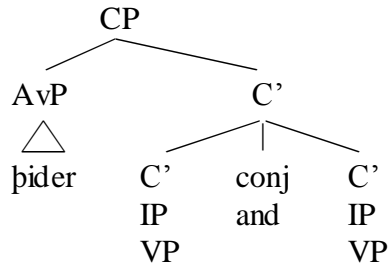
As before, we topicalize the Av to bring it to a position where it can be mistaken for a Cz.

15. þider I go and þider you go (þider is Av)



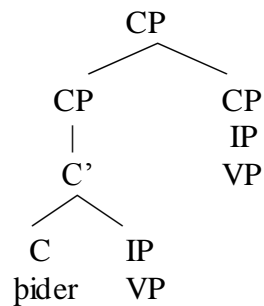
We then simplify the sentence, making the structure less strictly parallel. (If we kept it more parallel, as in the classic OE “þa ...þa” structure, one would be a Cz and the other would remain an Av.)

16. $\text{pider I go and you go}$ (pider is still Av)



We now reduce and remove the conjunction. This essentially leaves two non-coordinated verb phrases and the preceding adverb, which is now open to analysis as a Cz.

17. $\text{pider I go, you go}$ (pider is now a C)



3.9 A Problem with the Model

As I mentioned above, the meaning of the sentence is sometimes completely changed (in some cases reversed) over the steps of the model. “Before/after/without/until X object, Y occurred” is changed into “before/after/without/until Y occurred, [Z occurred].” A sentence with a complementizer must have multiple verbs, but a sentence without a complementizer can only have one (unless the verbs are coordinated). Where does the second verb phrase come from? This problem does not seem to occur in the $\text{Av} \rightarrow \text{Cz}$ model with the coordinate verb phrases.

This problem seems to pop up in the reduction stage: the elimination of the NP/DP requires that the ambiguous word gain a new argument, which is now a VP. It is this that changes the meaning of the sentence. If the reduction were less thorough, the sentence would not greatly change its meaning: “before X event, Y occurred” is changed into “before X event occurs, Y occurs,” or “without X object, Y occurs” is changed into “unless one has X object, Y occurs.” However, this does not work for every object or event X, without bringing in more concrete verbs. Simple verbs of possession or existence may be assumed to be understood. But for examples such as “after king X, king Y ruled” changing into “after king X died, king Y ruled,” a relatively concrete verb must be brought in, which was not assumed before.

Furthermore, if reduction to a pronoun does not occur, one is less likely to get the pronoun p-word that can be easily mistaken for a complementizer. Without the word of interest being able to “borrow” complementizer-ness by association and adjacency, I feel that the necessary reanalysis is highly unlikely. It does become somewhat more likely if the verb can be assumed or understood. “Until event X” and “until event X occurred” are not so very different semantically, but there is still a very large syntactic difference that I do not think can be overcome in only one reanalysis event.

3.10 Summary

Grammaticalization is the process by which a word loses lexicality in exchange for increased grammaticality. This occurs via a slow process of

relabeling and reanalysis over many generations of speakers, so that very small differences in interpretation and grammaticality can add up to noticeable and even large changes in a language. I propose a step-by-step model by which a preposition may grammaticalize into an adverb and thence into a complementizer. This process depends on the speaker's ability to treat a word that was formerly a preposition as part of a complex complementizer, thereby transferring the function from a pre-existing complementizer to a new one.

CHAPTER 4

QUANTITATIVE ANALYSIS OF THE TEXTS

4.0 Introduction

Chapter Four presents my data. The tables given here are condensed from the raw data in the appendices. I determine how many times each word appears as a given PoS, in what form: whether it is alone, in a phrase, or with accompanying accessory words, and how many and which type of accessory words. I note whether any particular part of speech tends to have more or fewer accessory words, and whether any author or text has noticeable patterns of usage.

4.1 Notes on the Data

There were a total of 2593 occurrences of my six words of interest in the texts. This includes a wide range of spelling variations, as well as occurrences of the words in conjunction with other “accessory words”. By accessory words, I mean words with which they co-occur and co-function. For instance, “oð to” is a complex preposition meaning “up to;” “to” is an accessory word to the word of interest, “oð.” As another example, “æfter þon þa þe” is a complex complementizer: the parts may be individually translated as “after then when that,” or they may be translated together as “after” or “when;” in either case I consider all the words beginning with þ to be accessory words to “æfter.” In the summary charts in this chapter, I have divided the accessory words into þ-words and non-þ words; the latter I abbreviate as acc. (By þ-words I mean any small grammatical word beginning with þ, generally þa, þe, or some derivative of se or

þæt.) I have divided them thus to better correlate the data to my model, since more-grammatical þ-words behave differently than less-grammatical accessory words in my model. Accessory words may be various parts of speech: þ-words are usually þa (Av or Cz), þe (Cz), þæt (Cz), or a form of se or þes (Pro). Acc-words are usually conjunctions (ah), prepositions (in, to), adverbs (sona), or complementizers (gif). When counting combined words such as “miððy” or “æfterðon,” I treat them as two words, “mið ðy” and “æfter ðon.”

Appendix A contains the full results of my data. For each word I give the number of occurrences as each part of speech. For adverbs I often break them down by meaning or translation, and sometimes give citations. For complementizers I give citations for most and translations or meanings for some. For most complementizers and adverbs I give their position in relation to the verb phrases in the sentence, to help verify their part of speech.

Appendix B contains a summary of my results, sorted by spelling variations of each form, as well as variations and combinations of accessory words and vel variants. (Vel, the Latin word for “or” was often used by the glossators to give multiple OE synonyms for Latin words or phrases which might be unfamiliar; e.g. “oððæt vel ða huil” tells the reader that the Latin word in question, “donec,” can mean either “until” or “while.”)

4.2 Æfter

Table 1

Æfter and its Accessory Words

Form		Farman			Owun			Lindisfarne		
		P	Av	Cz	P	Av	Cz	P	Av	Cz
Bare	æfter	24	15		60	4		75	2	
+ þ			3			9	11		30	15
+ 2þ			1			1	1		5	3
+ 3þ				2						
+ 4þ				2						
+ acc									2	
+ acc + þ							1			
+ 2acc + þ										
Totals		24	19	4	60	14	13	75	39	18

The bare forms of the word are always prepositions. This is nearly always true for all my words of interest, because I have made the editorial decision that demonstrative pronouns are accessory words rather than objects of prepositions, i.e. that “æfter ðas” (after these [things]) is an adverbial phrase (translatable as “afterwards”) rather than a prepositional one. This is in contrast to the occasions in which the demonstratives are determiners rather than pronouns, e.g. “æfter ðæm dagum” (after that day); I consider this an example of the bare form of the preposition.

Æfter and its variants (including spelling variants æfter, aefter, efter, æft, and æfterr) occur 159 times as prepositions, 72 times as adverbs, and 35 times as Cz. The adverbial uses of æfter can occur with or without an accessory word; nearly all of the accessory words are þ-words. Similarly, all the occurrences as

Czs contain at least one þ-word, and may contain as many as four. However, Czs never occur as bare forms without any accessories.

4.3 Ær

Table 2

Ær and its Accessory Words

Form		Farman			Owun			Lindisfarne		
		P	Av	Cz	P	Av	Cz	P	Av	Cz
Bare	ær	4	11		10	5		19	4	1
+ þ			1	10			8			7
+ 2þ				1			3			1
+ 3þ										
+ 4þ										
+ acc							2			1
+ acc + þ										
+ 2acc + þ										
Totals		4	12	11	10	5	13	19	4	10

Ær and its variants occur 33 times as prepositions, 21 times as adverbs, and 34 times as Czs. As with æfter, the prepositions are always bare forms. Here, however, the adverbs are nearly all bare forms as well; only one has an accompanying þ-word. The Czs nearly all have accessory words, most of which are þ-words, but one occurs as a bare form.

4.4 Butan

Table 3

Butan and its Accessory Words

Form		Farman			Owun			Lindisfarne		
		P	Av	Cz	P	Av	Cz	P	Av	Cz
Bare	butan	10	1		45	4	17	74	16	22
+ þ							1			2
+ 2þ										
+ 3þ										
+ 4þ										
+ acc						1	2			1
+ acc + þ							1			1
+ 2acc + þ										
Totals		10	1	0	45	5	21	74	16	26

Butan and its variants (including spelling variants buta, bute, butu, and butun) occur 129 times as prepositions, 22 times as adverbs, and 47 times as Czs. Again, the prepositions are always bare forms (one occurrence counted in the prepositions is actually a conjunction; see section 4.6). Nearly all of the adverbs are bare forms also, and the one exception has a non-þ accessory word. The Czs are mostly bare forms, but also occur with þ and non-þ accessory words.

4.5 Mid

Table 4

Mid and its Accessory Words

Form		Farman			Owun			Lindisfarne		
		P	Av	Cz	P	Av	Cz	P	Av	Cz
Bare	mid	84		1	285		4	492		3
+ þ				2	9	3	273	1	1	433
+ 2þ							21			38
+ 3þ							1			1
+ 4þ										
+ acc							3			3
+ acc + þ							13		1	19
+ 2acc + þ										
Totals		84	0	3	294	3	315	493	2	497

Mid and its variants (including spelling variants *mið* and *mip*) occur 871 times as prepositions, 5 times as adverbs, and 815 times as Cz. The prepositions are nearly all bare forms (the Lindisfarne exception is a case of over-literal translation and not really even a preposition, and the R1 exceptions are all occurrences where the scribe obviously meant to put “*mið*” rather than “*miððy*”). Farman uses *mid* rarely compared to the other two texts, and never as an adverb. The few uses of *mid* as an adverb all contain at least a þ-word. The few occurrences of Cz as bare forms are all, similar to the prepositions, places where the scribes meant to put “*miððy*” rather than “*mið.*” Otherwise all occurrences as Cz have at least one accessory word, nearly all of them þ-words, and often more than one.

The reason for such high numbers for this word is that the Lindisfarne glossator, and therefore Owun, use *mið*/*miððy* to translate the Latin word “*cum.*”

The Latin word can function as a preposition indicating accompaniment, in which case the glossators use “mið,” or as a Cz with a general meaning of “when,” in which case the glossators use “miððy.” “Miððy” is also used to indicate a tense change in the Latin text, since Latin has a finer gradation of tenses than OE, either actual tense change in a finite verb, or use of verbal participles. “Cum” is very common as both a P and a Cz in Latin, and therefore occurs very frequently in the OE glosses, sometimes several times in a single verse.

4.6 Oð

Table 5

Oð and its Accessory Words

Form		Farman			Owun			Lindisfarne		
		P	Av	Cz	P	Av	Cz	P	Av	Cz
Bare	oð	19 (2)						9		1
+ þ		1 (52)	2	12	2 (3)		18	2 (7)		21
+ 2þ										
+ 3þ										
+ 4þ										
+ acc			9			6			11	
+ acc + þ						6	2			1
+ 2acc + þ										
Totals		20 (54)	11	12	2 (3)	12	20	11 (7)	11	23

Oð and its variants (including spelling variants oþ, oþþ, and oðð) occur 97 times as prepositions (or conjunctions), 34 times as adverbs, and 55 times as Czs. This word poses some problems, because oððe (and other spelling variants, always with at least two þ/ð, and usually a single final vowel) is also a coordinating and correlative conjunction meaning “or.” I consider that their spelling indicates them to be some variation of oð + ðe, and therefore containing a

þ-word (except oðð, which since it lacks a final vowel, I count as a bare form). However, they are neither prepositions, adverbs, nor Czs, so I count them in the P column and give their numbers separately in parentheses. With these exceptions noted, the prepositions are mostly bare forms, although with a higher rate of exceptions than the other words of interest. The adverbial occurrences all have at least one accessory word, though most of them are non-þ-words. Most of these are modifiers of prepositions, which could also be considered complex prepositions (I could have counted them under preposition +acc, but chose not to), usually meaning “up to” or “into.” The Cz occurrences, excepting one bare form, all have at least a þ-word accompanying them.

4.7 Wið

Table 6

Wið and its Accessory Words

Form		Farman			Owun			Lindisfarne		
		P	Av	Cz	P	Av	Cz	P	Av	Cz
Bare	wið	34			23		1	57	4	15
+ þ										2
+ 2þ										
+ 3þ										
+ 4þ										
+ acc									22	
+ acc + þ							1			3
+ 2acc + þ										2
Totals		34	0	0	23	0	2	57	26	22

Wið and its variants (including spelling variations wiþ, wiðe, and uið)

occur 114 times as prepositions, 26 times as adverbs, and 24 times as Czs.

Prepositions occur only in the bare form, and Farman never uses it as any other

part of speech. Only the Lindisfarne glossator uses it as an adverb, and most of these (all the +acc occurrences) are modifiers of prepositions, as with “oð.” Unusually, most of the Cz occurrences are bare forms, but also occur with þ and non-þ accessory words.

4.8 Summary

I have a total of 1338 prepositions, 180 adverbs, 1010 Czs, and 65 conjunctions. Most of the prepositions are bare forms but that is mostly because of how I decided whether or not they are prepositions. Many of the adverbs could actually be counted as compound prepositions. Most acc-words are þ-words, and are mostly in Czs, and can occur in combinations of up to four acc-words. The number of accessory words is inversely proportional to frequency. Ær, butan, and oð never have more than two accessory words; wið and mið may have up to three; only æfter ever has four. There are 1457 bare forms, 1004 occurrences with one accessory word, 124 with two, 6 with three, and 2 occurrences with four accessory words. Farman has 303 occurrences over one gospel, Owun 860 over three gospels, and Lindisfarne 1430 over four.

CHAPTER 5

THE MODEL APPLIED TO THE DATA

5.0 Introduction

In Chapter Five I analyze my data by applying the words to my model from Chapter Three to see whether the data supports each step. Rather than giving repetitive examples and diagrams for each word in each step of the theory (some of the words do not have occurrences at every step), I have chosen to give examples and diagrams for only one word, *æfter*, which has occurrences in nearly every step.

5.1 Non-initial Prepositional Phrase

In the first step of the model (example 8), a sentence contains a non-initial prepositional phrase. Here the PP “*æfter dagum*” appears at the very end of the sentence.

18. *& eftersona infoerde capharnaum ða burg æfter dagum*

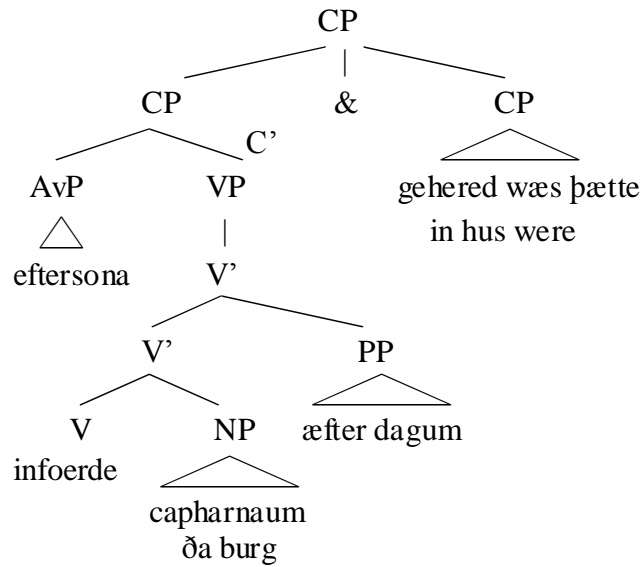
 & soonafter went.in Capernaum the city after days

 & gehered wæs þætte in hus were

 & heard it.was that in house he.was

 “And soon after some days, he went into the city Capernaum, and it was heard that he was in the house.” (Lindisfarne Mk 2.1)

19.



5.2 Initial Prepositional Phrase

In the next step of the model (example 9), the PP is moved by topicalization to the beginning of the sentence. This is a somewhat complex PP, since the translator gives synonyms for the preposition itself, but the synonyms can be translated as a single preposition, which still has an NP or DP object rather than a pronoun.

20. *& æfter vel ymb lytle huile geneolecdon ða ðe stodon*

& after or around little while came.near they who they.stood

& cuoedon to petre soðlice ðu of ðæm arð forðon

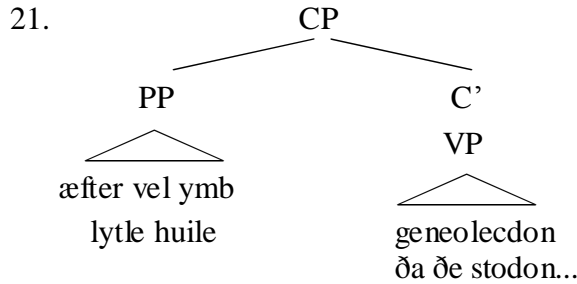
& they.said to Peter truly you of them are because

& reord ðin cuð vel cyðic ðec doeð

& speech your known or aware you makes

“And after a little while the bystanders came near and said to Peter, ‘Truly you are one of them, because even your speech makes you known.’”

(Lindisfarne Mt 26.73)

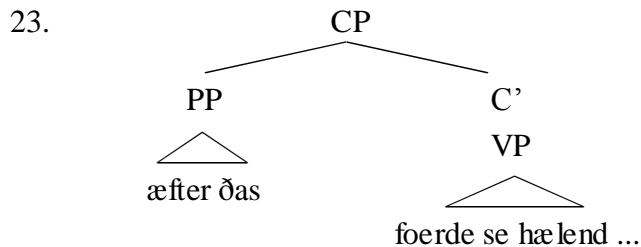


5.3 Reduced Prepositional Phrase

In the next step (example 11), the object of the preposition is reduced from a full NP or DP to a simple pronoun, “ðas.” (In this example the glossator has chosen to substitute “þæt luh” for the name “Tiberiadis;” glossators often did not translate proper nouns.)

22. *æfter ðas foerde se hælend ofer sae galiles*
 after these [things] went the healer over sea of Galilee
þæt is þæt luh
 which is that lake

“After this Jesus went over the Sea of Galilee which is also called Tiberias.” (Lindisfarne J 6.1)



5.4 Complex Complementizer

In this step (example 12), the former preposition gradually loses its accessory words and can be reanalyzed as an adverb or a complementizer. This step is the reanalysis, and depends on the accessory words: a word such as *þa*, which a speaker can easily believe to be either a Cz or a pronoun, is more easily reanalyzed than a word such as *þæm*, which is not easily identified as a Cz, and is therefore likely to either require more effort and time to reanalyze, or to stop at the intermediate step of adverbial phrase.

23. *& æfter ðon gefylled wer dagas clænsunges his æfter ae*

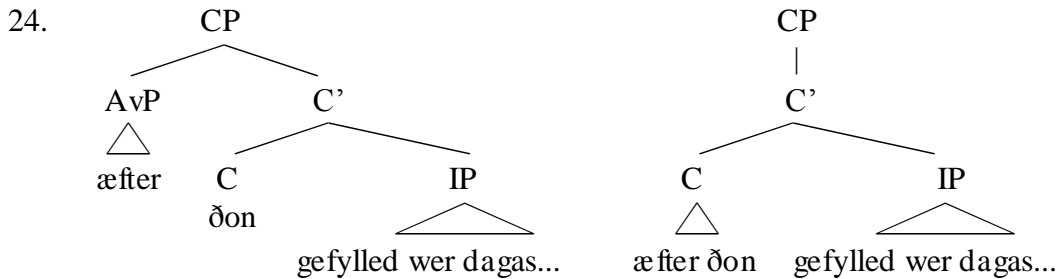
& after when fulfilled were days of.cleansing his according.to law

lædon hine in hierusalem

they.led him into Jerusalem

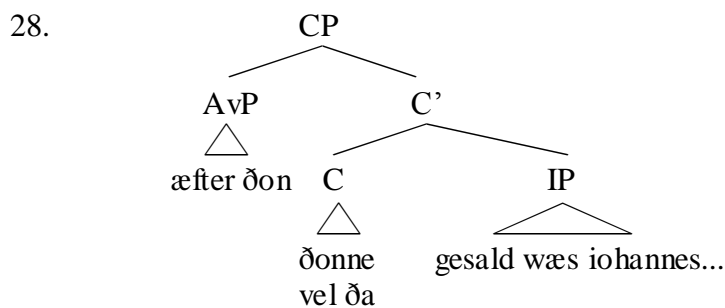
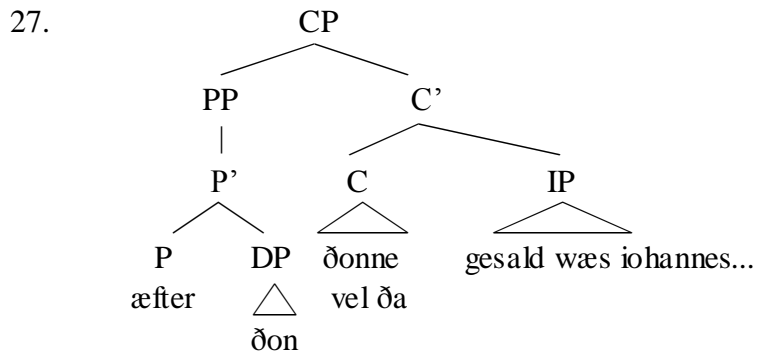
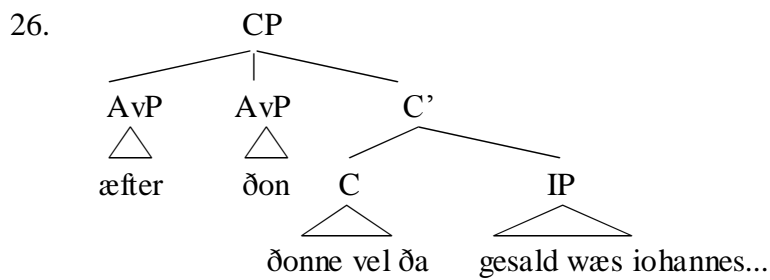
“And after the days of purification were fulfilled according to the law,

they brought him into Jerusalem.” (Lindisfarne L 2.21)

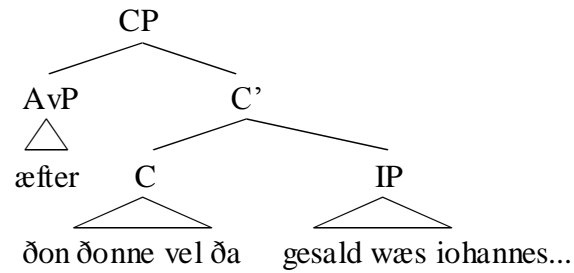


In this second example, we have a sentence with multiple *þ*-words which need to be reduced and reanalyzed. They can be reanalyzed in various combinations or drop off in various orders, leading to multiple paths of reanalysis.

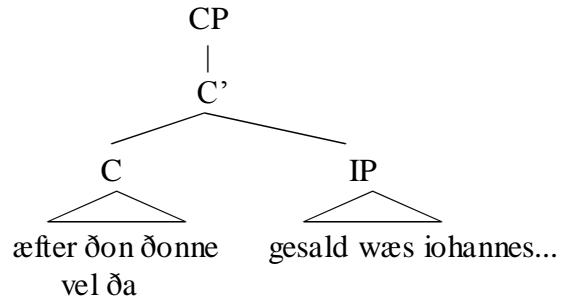
25. *æfter ðon ðonne vel ða gesald wæs iohannes cuom se hælend*
 after then when or when sold was John came the healer
in galilea bodade godspell rices gode
 into Galilee announced gospel of.kingdom God
 “After John was betrayed, Jesus came into Galilee preaching the gospel of
 the Kingdom of God.” (Lindisfarne Mk 1.14)



29.



30.



5.5 Simple Complementizer

This is the final step of the model (example 13). Unfortunately “æfter” has not grammaticalized all the way: it is not available in bare form as a Cz. I will use “ær” for this example instead, even though it has probably not grammaticalized all the way either: this sentence is the only example of a bare form Cz for “ær,” and is a repetition of an earlier verse, J 13.19, which uses ær ðon.

31. & nu ic cuoed iuh ær ðæt sie þætte miððy auorden se

& now I tell you before that is that when becomes is

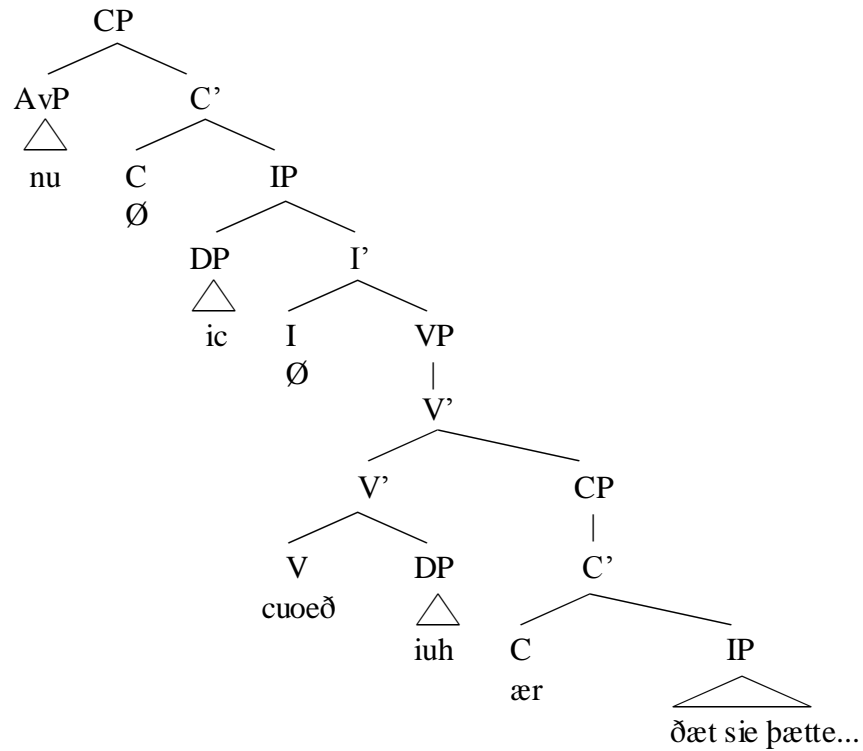
vel bið gie gelefe

or may.be you believe

“And I tell you before it happens, so when it happens you will believe.”

(Lindisfarne J 14.29)

32.

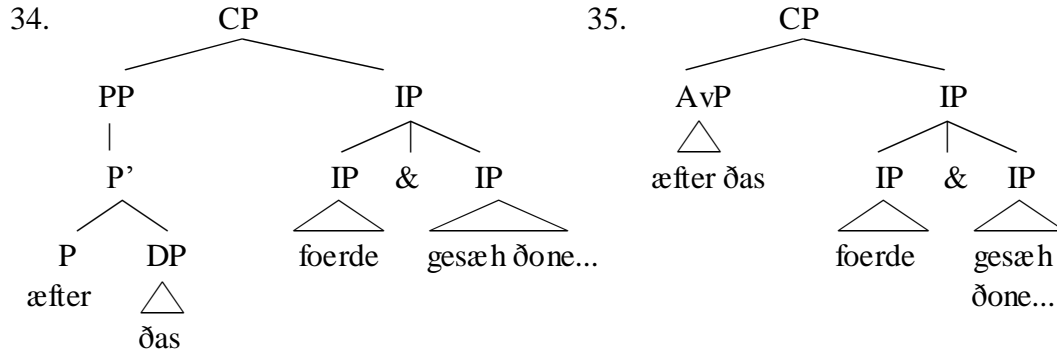


5.6 Adverbs

Any initial PP can have adverbial function, as above (examples 21 and 23). We can also have a bare word or word+acc as an adverb if there is only one VP involved. In this first example, a reduced prepositional phrase is reanalyzed as an adverbial phrase. (In this sentence also, the glossator does not translate a name and the word “telonium” [custom-house], which he must have considered either untranslatable or familiar to the audience as a Latin word.)

33. *& æfter ðas* *foerde & gesæh ðone bærsynnig* *genemned*
 & after these [things] went & saw the tax-collector named [Levi]
 wæs sittende to *& cuoeð him* *fylg* *mec vel soec mec.*
 was sitting at [custom-house] & said to.him follow me or seek me.

“And afterwards he went and saw the tax collector called Levi sitting at the custom-house and said to him, ‘Follow me.’” (Lindisfarne L 5.27)

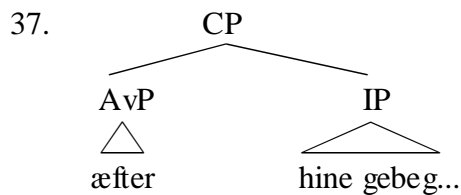


In this second example, the bare form is used as an adverb. (The pronoun hine is the direct object of the reflexive verb gebeg, rather than the object of the preposition; note that it is accusative rather than dative.)

36. *& æfter hine gebeg aurat on eorðu*

& again him bow wrote on earth

“And bending over again, he wrote on the ground.”(Lindisfarne J 8.8)



5.7 Summary

Although not every step of my theory has examples of every word, each step has examples from multiple words, and each word can occur at multiple steps. For a word to occur at all the steps of my theory, it should occur as a bare form in all three parts of speech, and as both adverb and complementizer with at least one accessory or þ-word. This is true only of ær (but see note in section 5.5).

However, most words have occurrences in nearly all of these groups, the most common gap being in the intermediate adverb stage, and there is certainly enough information between all six words to create a complete picture. All the stages of my model are represented within the examples, showing that my model fits with the patterns of actual language use.

CHAPTER 6

ANALYSIS AND DISCUSSION

6.0 Introduction

In this chapter I assess where each word is on a grammaticalization cline of $P \rightarrow A \rightarrow Cz$, and determine whether any particular text or dialect shows evidence of more grammaticalization than another. I go word by word and compare how often each is used with and without single or multiple accessory words, and whether they are used differently in the different texts. Finally, I compare the patterns of the texts themselves to determine whether any one shows a more advanced stage of grammaticalization.

6.1 *Æfter*

As seen in the previous chapter, *æfter* as a bare form can function as a P or Av, but not Cz. All instances of *æfter* as a Cz involve a β -word, rather than a non- β accessory word (one occurrence has a non- β , but only co-occurring with a β -word). *Æfter* can function as an adverb either alone, or with a β or non- β accessory word. Owun and Lindisfarne use *æfter* similarly, but Farman does not use it with any non- β accessory words, and only uses it as a complementizer with three or four accessories, while Owun and Lindisfarne use only one or two accessories.

Cline: it is thoroughly P, pretty well established as Av, but only Cz with the assistance of grammatical words.

6.2 Ær

This word is used very much like *æfter*. The bare form occurs often as a P or Av, but only once as a Cz (see note about example in chapter 5). Most instances of *ær* as a Cz involve a β -word. The adverbial uses are usually bare forms, with only one exception by Farman, using a β accessory word. Owun and Lindisfarne use it as a Cz with either a β or non- β accessory word, but never both; Farman uses only β accessories.

Cline: it is thoroughly P and Av, but only Cz with the assistance of grammatical words.

6.3 Butan

The bare form of *butan* can occur as any part of speech, and in fact most instances of Cz are a bare form. It never uses more than one β -word, but may use a β and non- β together. Farman uses it only as P and Av, never Cz. Almost all the adverbial uses are the bare form, and the only exception uses a non- β accessory word. As a Cz it usually has the bare form, but may also have β , non- β , or a combination of accessory words.

Cline: Thoroughly P and Av, and pretty well along as Cz in Owun and Lindisfarne.

6.4 Mid

Mid occurs in its bare form nearly always as a P. It rarely occurs as a Cz, but these are certainly all errors for *miððy*. Interestingly, the bare form is never an Av; Farman never uses it as an Av at all. Accompanied by a β -word, it sometimes

occurs as a P or Av, and very often as a Cz. Most of the Czs are able to function with only a single þ-word as the only accessory, but they often need multiple þ-words, a combination of þ and non-þ, and only rarely have a non-þ accessory word alone.

Cline: Thoroughly P, but in Lindisfarne and Owun seems to have mostly skipped Av and gone straight to a strong showing as Cz, though always accompanied by grammatical words.

6.5 Oð

Oð has a unique behavior in that, when in combination with a single þ-word, it is usually a conjunction. The conjunction form is certainly related to the others, but seems to be on a different branch of the Grammaticalization cline. The bare form is usually a preposition, rarely a conj or Cz. When an Av, it always has some sort of accessory word, but these are mainly non-þ words (these instances are closer to complex prepositions than oð + acc). As a Cz, it always has a þ-word and no non-þ words, with the exception of one bare form.

Cline: a branching one, I think. One version starts at P and goes directly to a coordinating/correlative conjunction. The other branch is pretty strongly a P, but can only work as Av or Cz with the assistance of grammatical words; it has a stronger showing as a Cz than as Av.

6.6 Wið

Wið occurs in its bare form usually as a preposition, but Lindisfarne often uses it as a Av or Cz. Farman uses it only in its bare form and only as a P. Owun

uses it nearly the same way, except once as a bare Cz and once as a Cz with two accessories. Lindisfarne uses it most often as a bare P, but also uses it as an Av, either bare or more often with a non-þ accessory (like oð, these are closer to complex prepositions). Lindisfarne also often uses wið as a Cz, either bare or with at least one þ-word accessory; it also twice uses it with three accessories: one þ and two non-þ, the only occurrences of this type.

Cline: Very strongly P, rather weakly Av, and a pretty decent showing for Cz, but only in Lindisfarne.

6.7 Differences Between Texts

Farman is not fond of using these words as complementizers, not nearly as much as Owun and Lindisfarne. He uses them almost 60% of the time as prepositions, and the rest divided about equally between Conj, Av, and Cz. He uses the words in question less often in general (.77% of the text), and only .00075% of his text is these words as Czs. He must use Czs, since one cannot make useful sentences without them, but they must be ones not investigated in this project.

Owun and Lindisfarne behave similarly to each other, percentage-wise, as is to be expected. They use the words in question as prepositions about 50% of the time, and about 40% of the time as Czs; Avs are relatively rare, and conjs extremely so. Most of this is due to the very high numbers of mið and miððy. Eliminating all variants of mid from Owun and Lindisfarne give them rates very

similar to Farman's (about 60% P, and approximately equal amounts of Av and Cz, though still very few Conj).

There is no general trend towards later texts using more grammaticalized forms; in fact the opposite is true. Farman uses fewer grammaticalized forms and less often. Even comparing Owun and Lindisfarne, Owun has a significantly lower Av/Cz ratio. One would expect Grammaticalization to increase over a span of 100 years. Though not by much: possibly grammaticalized versions of other words have out-competed the words I am looking at, or the people of Farman and Owun's time prefer Czs that cannot be confused for other parts of speech. Probably also much of the difference is due to style. For instance, Farman uses þa and þonne to translate Latin cum, but he also uses mid/mið as a P only two-thirds as often as Owun and Lindisfarne. There isn't really much a writer can use instead: he doesn't substitute ablative or some other construction. Perhaps some of it is actually the style of the Latin gospel writers: if Matthew (and his later editors) tended to say "Jesus and his disciples went" more often than he said "Jesus went with his disciples," then that will throw off Farman's rate no matter how often he uses grammaticalized forms in his everyday speech.

6.8 Trends and Conclusions

There are two possible sources of variation: diachronic and geographic. Both Farman and Owun, writing one hundred years after the Lindisfarne glossator, ought to show diachronic variation from Lindisfarne. One would expect the later texts to show a higher level of grammaticalization, since this process

proceeds with time and does not generally reverse itself. Farman, writing in his Mercian dialect different from the Northumbrian dialect common to Owun and Lindisfarne, ought to show geographic variation. One would expect the Northumbrian dialect, being in greater contact with ON speakers, to show greater advances in grammaticalization. However, the difference between Northumbrian and Mercian would be less than that between Northumbrian and West Saxon, and the difference between south Northumbrian and north Mercian is likely to be small indeed.

Contrary to diachronic expectation, Farman is a very conservative syntactician. His uses of the words in question are mainly as prepositions, and much less often as adverbs or complementizers. As I stated before, it is nearly impossible to translate or gloss large amounts of coherent sentences without some sort of complementizer. And yet he uses these words less often than either Owun or the Lindisfarne glossator, as any part of speech. It is likely that the Cz and other forms Farman preferred were not the ones I chose to investigate. For instance, I did not look at any wh- words: a form of “when” could be used for *miððy*, and “while” for derivatives of *oð* and *wið*. Likewise, if he preferred the old-fashioned OE *þa* as his complementizer of choice, I would not have noticed.

However, given that he must use some sort of Cz, it is interesting that the ones he uses do not seem to be of the type I am investigating. Perhaps he finds the grammaticalized forms to be incorrect or ungrammatical for his dialect. It seems

that either Mercian in general or Farman in particular is very conservative about using former prepositions as adverbs and complementizers.

Owun uses grammaticalized forms in about the same proportions as his Lindisfarne original. He uses prepositions almost as often as complementizers, but adverbs rarely. He also uses them with about the same frequency: for both Owun and Lindisfarne, the words in question occupy just under 1% of the text, while Farman only has about 0.75%. Like Farman, he probably used other forms that I did not happen to investigate. Owun seems to be midway the two extremes of Farman and Lindisfarne. This is not surprising given that he is geographically similar to Lindisfarne and diachronically similar to Farman.

The Lindisfarne glossator seems to be very innovative in his use of grammaticalized forms. Since Owun copied Lindisfarne rather closely, much of Owun's usage can be attributed to Lindisfarne's example. Likewise, the Lindisfarne gospels have rather interesting and unique patterns of spelling and inflectional endings, which Owun tends to regularize to something more closely approaching the West Saxon standard. Since Lindisfarne is the main example of the Northumbrian dialect in the OE corpus, it is difficult to tell whether these innovations in spelling or syntax are common to the whole dialect, or only this glossator. But certainly Owun's greater use of grammaticalized forms than Farman is directly copied from Lindisfarne, rather than being any sort of dialectal advancement. It is interesting that Owun found the spelling odd enough to require extensive correction, but not the grammatical innovations. This likely means that

they felt more correct to him than they did to Farman, who chose to change the syntax as well as the spelling.

As one would expect, the different words grammaticalize at different rates. However, none of the words stand out as very much more or less grammaticalized than any other. (Although *mið/miððy* has a disproportionately high frequency, its ratio of grammaticalized to ungrammaticalized forms is unexceptional.) While Farman tends to use only the less grammaticalized versions of each word, Owun does not seem to have objected to the Lindisfarne glossator's decisions of correctness. There are no words for which Owun and Lindisfarne disagree on the degree of grammaticalization or position on the $P \rightarrow Av \rightarrow Cz$ cline. Farman's greatest disagreement is his insistence on using *oð* mainly as a conjunction, thereby forming a branch on that particular cline.

Sadly, I am not able to detect any sweeping trends in grammaticalization, either geographically or diachronically. The different patterns of usage seem to be due mainly to authorial and editorial individuality. However, my step-by-step model is well populated by enough examples to give it empirical as well as theoretical soundness.

REFERENCES

- Alfred, King. King Alfred's West-Saxon Version of Gregory's Pastoral Care. Ed. Henry Sweet. EETS 45, 50. London: Oxford UP, 1871.
- "The Anglo-Saxon Chronicle." (Parker MS; Corpus Christ College, Cambridge MS 173 ff. 1-32). The Labyrinth Project. 1994. Georgetown University. 15 October 2010. <<http://www8.georgetown.edu/departments/medieval/labyrinth/library/oe/texts/asc/a.html>>
- Campbell, Alistair. Old English Grammar. Oxford: Clarendon P, 1959.
- Fischer, Olga, Ans van Kemenade, Willem Koopman, and Wim van der Wurff. The Syntax of Early English. Cambridge: Cambridge UP, 2000.
- Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie, eds. World Atlas of Language Structures. Oxford: Oxford UP, 2005.
- Heine, Bernd and Tania Kuteva. Language Contact and Grammatical Change. Cambridge Approaches to Lang. Contact 3. Cambridge: Cambridge UP, 2005.
- Hopper, Paul J. and Elizabeth Closs Traugott. Grammaticalization. Cambridge: Cambridge UP, 2003.
- Kiparsky, Paul. "Indo-European Origins of Germanic Syntax." Clause Structure and Language Change. Eds. Ian Roberts and Adrian Battye. Oxford, Oxford UP, 1994. 140-167.
- Mitchell, Bruce, and Fred C. Robinson. A Guide to Old English. 6th Ed. Malden, MA: Blackwell, 2001.
- Mossé, Fernand. A Handbook of Middle English. Trans. James A. Walker. Baltimore: Johns Hopkins P, 1968.
- Pintzuk, Susan, George Tsoulas, and Anthony Warner. "Syntactic Change: Theory and Method." Diachronic Syntax: Models and Mechanisms. Eds. Susan Pintzuk, George Tsoulas, and Anthony Warner. Oxford: Oxford UP, 2000. 1-22.
- Roberts, Ian. "Grammaticalization, the Clausal Hierarchy and Semantic Bleaching." Gradience, Gradualness and Grammaticalization. eds. Elizabeth Closs Traugott and Graeme Trousdale. Typological Stud. in Lang. 90. Amsterdam: Benjamins, 2010. 45-73.

- Santorini, Beatrice, and Anthony Kroch. The Syntax of Natural Language: An Online Introduction Using the Trees Program. 2007.
 <<http://www.ling.upenn.edu/~beatrice/syntax-textbook>> October 2010.
- Skeat, Walter W., ed. The Gospel According to Saint Matthew and According to Saint Mark. Darmstadt: Wissenschaftliche Buchgesellschaft, 1970.¹
- Skeat, Walter W., ed. The Gospel According to Saint Luke and According to Saint John. Darmstadt: Wissenschaftliche Buchgesellschaft, 1970.¹
- Toon, Thomas E. The Politics of Early Old English Sound Change Quantitative Anal. of Ling. Structure 2. New York: Academic P, 1983.
- Traugott, Elizabeth Closs. The History of English Syntax: A Transformational Approach to the History of English Sentence Structure. Transatlantic Ser. in Ling. 7. New York: Holt, 1972.
- van Gelderen, Elly. Grammaticalization as Economy. Linguistics Today 71. Amsterdam: Benjamins, 2004.
- Whitman, John. "Relabelling." Diachronic Syntax: Models and Mechanisms. Eds. Susan Pintzuk, George Tsoulas, and Anthony Warner. Oxford: Oxford UP, 2000. 220-238.

¹ Citations reflect the fact that each Gospel (with its respective notes) is paginated separately.

APPENDIX A

TOTAL OCCURENCES BY WORD

Forms		Farman			Owun			Lindisfarne		
Bare	Final	P	Av	Cz	P	Av	Cz	P	Av	Cz
æfter	æfter	24	2		60	3		73	1	
	æfter þon		2							
	æfter þon þæt		1							
	æfter þon þe			2						
	æfter þon þa þe			1						
	æfter þon þanne þe			1						
	æfter ðas				3				9	
	æfter ðæm					1			1	
	æfter ðon ðone					1				
	æfter ðon					4	11		9	12
	æfter ðon ða						1			
	æfter ðisse								1	
	æfter ðis								1	
	æfter ðis ðona								1	
	æfter ða								3	
	æfter ðas ða								1	
	æfter ðas ðonne								2	
	ðona æfter ðas								1	
	æfter ðon ðe									1
	þæt ne æfter ðon									1
	æfter þætte								2	
	æfter							2		
	æfter ðon									1
	æft		13							
	æfterr								1	
	æfterðon								4	2
	gif æfterðon						1			
	æfterðon ðe									1
	efter					1				
	efter þa		1			1				
	efter sona								1	
	sona efter								1	
ær	ær	4	11		10	5		19	4	1
	ær þon		1							
	ær ðon						4			6
	buta ær						1			1
	ah ær						1			
	ær ðætte						1			
	ær ðon þæt									1
	ærðon			1			3			1

	ærðon ðonne						1			
	ærðon ðe						2			
	ærþon			9						
	þæt ærþon			1						
butan	butan	10	1		4			1		
	ah butan					1				
	buta				39	4	17	71	16	22
	buta mið ðy						1			
	buta ær						1			1
	ah buta						1			
	buta þætte									1
	buta ðætte						1			
	buta þæt									1
	buta miððy									1
	bute							2		
	butu				1					
	butun				1					
mid	mid	54			1			2		
	mid þy þonne þende						1			
	miþ ðy						1			
	mið	30		1	284		4	490		3
	mið ðy						90			3
	mið ðon								1	
	mið ðy ðonne						1			
	mið ðy ðæt						1			
	mið þy			1		3	6			
	miððy			1	9		176	1		430
	miððy ðonne						2			6
	miððy ðætte						1			
	miððy þæt									5
	miððy sona								1	
	miððy forðon						8			6
	miððy ðona									2
	ða miððy þæt									1
	ah miððy						3			3
	forðon miððy						5			9
	æfterðon miððy									1
	ða miððy						4			9
	ðonne miððy						5			7
	ðætte miððy						7			
	wið ðon miððy									1

	wið þæt miððy								2	
	þætte miððy								9	
oð	oð	6					4			
	oð to		2			5				
	oð on					1				
	feorra oð							1		
	oþ	13								
	oþ to		4							
	oþ þæt			5						
	oðð						5		1	
	oðð to							5		
	oðð on							4		
	oðð in							1		
	oðð ðæt								1	
	oðð ða hwil								1	
	oðð þætte								1	
	oþþ	2								
	oþþæ	1								
	oþþa	4								
	eþþa	2								
	oþðe	2			1					
	oþþe	1								
	oððæ					1	2			
	oþþæt			5						
	oþþætte			1						
	oþþætti			1						
	oðþæt								4	
	oðþætte								1	
	oððþæt								1	
	oððþætte								1	
	oððdæt								1	
	oþþe	43	2							
	oþþe to		3							
	oððe				4		5			
	oððe on					5				
	oððe ða hwyle						1			
	oðða							2		
	oðða ða hwile						1			
	oððæt						17		11	
	oððæt in					1				
wið	wið	30			23		1	51	4	15
	wið hwile ða						1			

	wið inn								1	
	wið on								3	
	wið to								16	
	wið ða huile									1
	wið ða hwile									1
	ðona wið									1
	wið þæt miððy									1
	wið ðon miððy									1
	ða huile wiðe									1
	wiððy									1
	wiþ	4								
	uið							6		
	uið to								2	

APPENDIX B
THE RAW DATA

Notes: Citations are given in chapter.verse format; all citations of Farman are the gospel of Matthew. Color code: P, Av, Cz, (Cj), *uncertain*. Occasional translations are given in Latin or MnE when relevant to the definition of the word, and the relation to VPs are noted for most Czs.

Word	Farman	Owun	Lindisfarne
æfter	24 + 5 + 4	60 + 12 + 12	72 + 30 + 15 + 2
	prep NP 24	prep NP 60	prep NP 72
	4:2 VP æfter þon VP	æfter ðas = post hæc 3 (L10.1, 12.4, 17.8)	J8.8 æfter VP
	16.21 æfter þon VP	Mk4.19 = adj/adv	J 5.14 æfter ðæ m VP
	21.9 æfter VP (behind)	J1.30 forðon vel æfter ðæ m	Mt26.64 heona vel æfter ðisse VP
	21:32, æfter þon þæt VP	J12.22 = back/again	J11.11 æfter ðis VP
	24.21 æfter VP (future)	J13.36 = after	J2.12 æfter ðis ðona VP
	27.31 æfter þon þe VP VP	Mk 4.8 æfter ðon ðone ... soðða = deinde ... deinde	J19.28, 21.1 æfter ða VP
	27.35 æfter þon þa þe VP VP	Adv æfter ðon = deinde/postea (Mk8.25, J5.14, 13.12, 20.27)	J13.7 VP æfter ða
	27.44 VP æfter þon þe VP	æfter ðon VP VP (Mk1.14, 14.28, 15.20, 16.19, L2.21, 2.22, 12.5, 14.29, 18.33, 23.33, 13.12)	L 5.27 VP æfter ðas
	26.32 æfter þon þanne þe VP VP	L22.20 (VP) æfter ðon ða VP	æfter ðas VP (L6.26, 12.4, 17.8, J3.22, 5.1, 6.1, 7.1, 19.38)
			æfter ðas ðonne VP (Mk16.12, L10.1)
			L18.4 æfter ðas ða VP
			J11.7 soðða vel ðona æfter ðas VP
			æfter ðon VP (Mt 4.2, 21.29, 26.16, L7.11, 8.1, 8.12, 16.7)

			VP æfter ðon (Mt21.32, J3.36)
			æfter ðon VP VP (Mt26.32, 27.31, Mk16.19, L2.21, 2.22, 12.5, 15.14, 15.30, 23.33)
			L20.33 VP æfter ðon VP
			Mt27.35 æfter ðon ðonne VP VP
			Mk1.14 æfter ðon ðonne vel ða VP VP
			L14.29 þætte ne æfter ðon VP VP
			L18.33 æfter ðon vel siðða VP VP
			L22.20 æfter ðon ðe VP VP
			L2.24, 22.22 æfter þætte VP = <i>secundum</i>
efer	1	2	2
	21:30, efter þa VP	J7.45 VP efter	Mk10.1 VP efter sona
		J13.7 VP efter þa	Mk14.72 & sona efter VP
æfter			1+2
			Mk15.20 & æfter ðon VP VP
			J19.5 prep = <i>secundum</i>
			J20.26 prep NP
æfterr			1
			Lk17.30 adv = <i>secundum</i>
æft	13		
	Adv = back/again/after 13		
æfterðon		1	7
		Lk20.33 VP gif æfterðon VP	Mk4.17 Adv = indeed
			Mk4.28 Adv = then
			Mk8.26 = then

			Mk14.28 æfterðon ðe VP VP
			J1.31 VP foreðon vel æfterðon VP
			J13.13 æfterðon vel forðon VP VP
			J20.27 Adv = then
ær	4 + 11+ 1	10 + 5 + 7	19 + 4 + 9
	prep NP 4	prep NP 10	prep NP 19
	Adv 11 (often for plup)	Adv 5 (L 23.12, J15.18, 19.41, 20.4, 21.1)	Adv 4
	6.8 VP ær þon VP	VP ær ðon VP (L2.21, J4.49, 13.19)	ær ðon VP VP (Mt 26.34, 26.75, Mk14.30)
		L2.26 VP buta ær VP	VP ær ðon VP (L22.15, J13.19)
		L21.12 ah ær VP VP	L2.26 VP buta ær VP
		J8.58 ær ðon VP VP	J1.48 ær ðon VP (miððy VP) VP
		J14.29 VP ær ðætte VP	J4.49 VP ær ðon þæt VP
			J14.29 VP ær VP (same sent as 13.19)
ærðon	1	6	1
	Mt17.9 VP ærðon VP	Mk14.31 VP ærðon ðonne VP	Lk22.61 ærðon VP VP (w/in that-clause)
		Mk14.72 ærðon ðe VP VP	
		Lk22.15 VP ærðon VP	
		Lk22.61 = Mk14.72	
		J1.48 ærðon VP (miððy VP) VP	
		J17.5 VP ærðon VP	
ærþon	10		
	1.18 VP ærþon VP		
	5.18 VP ærþon VP (before/until) (Mt5.18, 5.26,		

	10.23, 16.25, 24.34, 24.39)		
	23.29 VP þæt ærþon VP (before/until)		
	26.34 ærþon VP VP (Mt26.34, 26.75)		
butan	10 + 1	4+1	1
	butan NP 10	butan NP 4	butan NP
	23.26 = outside	1.45 ah butan (PP) wære	
buta		39 + 4 + 21	70 + 16 + (1) + 26
		prep NP 39	prep NP 70
		Adv 4 (Mk 14.68, L13.28, J16.2, 18.16)= out	Adv = out 5
		Mk9.9 VP buta mið ðy VP	Adv = only 5
		L2.26 VP buta ær VP	Adv+PP 6
		L13.3 ah buta VP VP	(1 conj Mk6.5)
		J10.10 VP buta ðætte VP	Mk 9.9 VP buta miððy VP
		VP buta VP (Mk 6.5, 7.3, L9.13, J1.18, 3.2, 6.44, 6.65, 7.51, 15.4, 15.4, 19.11)	L2.26 VP buta ær VP
		buta VP VP (Mk13.20, J3.3, 4.48, 6.53, 9.33, 20.25)	J10.10 VP buta þætte VP (nisi ut)
		(one result is a typo/mistake)	Mt5.13 VP buta þæt VP (nisi ut)
			buta VP VP (Mt5.20, 18.3, 24.22, Mk7.3, 13.20, L13.3, J3.3, 3.5, 4.48, 6.53, 9.33, 12.24, 20.25)
			VP buta VP (Mt12.29, L9.3, J3.2, 6.44, 6.65 7.51, 15.4, 15.4, 19.11)

bute			2
			J18.16, 19.13 Adv = out
butu		1	
		Mk12.8 prep NP = out of	
butun		1	
		Mk11.13 nowiht butun NP	
mid	54	1	2
	= cum 46	prep NP	prep NP 2
	= apud 3		
	= ablative 5		
miþ ðy		1	
		Mk1.37 & miþ ðy VP VP	
mid þy	1		
	26.6 mid þy þonne þende VP VP		
miððy	1	211 + 9	491 + 1 + 1
	J18.1 miððy VP VP	(none in Mk)	= cum 78
		= cum 81	= dum 5
		= dum 15	= abl 9
		= con 2	= tense 12
		= quoadusque 1	Mt13.5 miððy sona
		= other/tense 26	= cum 77
		= cum 82	= dum 1, deinde 1
		= dum 1	= abl 23
		= cum (for mid) 3	= tense 14
		= abl abs 3	= quemcumque 1
		= other 6 (all with cum	= cum 110

		somewhere)	
			= dum 17
			= ut 3
			= abl 24
			= tense 13
			= cum 88
			= dum 1
			= ut 3
			= abl 7
			= tense 5
mið	30 + 1	284 + 4	490 + 3
	= cum 23	= cum 59	= cum 71
	= apud 5	= ablative 40	= apud 8
	= ex 1	= aput 6	= abl 38
	prep NP = 1	= in 1	Mt13.21 mið ðon vel sona
	= cz (for miððy) J18.3	=? (Mk6.37) 1	= cum 62
		= cz (for miððy) 1	= apud 6
		= cum 64	= abl 46
		= ablative 27	= simul 2
		= apud 5, aput 3	= cz 2 (prob for miððy)
		= cz 3 (for miððy)	= cum 82
		= other/weird 9	= apud 12
		= cum 41	= abl 69
		= apud 5, aput 7	= ex, in, prae, 3 (one each)
		= ablative 17 (usu instr)	= simul 2
			= weird 1
			= cum 50

			= apud 13
			= abl 24
			= <i>weird 1</i>
			= cz 1 (prob for miððy)
mið ðy		92	3
		= cum 65 (1 --ðonne, 1 -- ðæt)	Mt5.11 VP mið ðy VP
		= deinde 1, dum 2	Mt12.43 mið ðy VP VP
		= other (ablabs, pspt, etc) 24	Mt15.19 & mið ðy VP VP
mið þy	1	3 + 6	
	25.31 & mið þy VP & VP	Adv = then 3	
		VP mið þy VP VP (Mk1.32, 2.15)	
		& mið þy VP VP (Mk 1.42, 2.4, 2.14)	
		Mk2.5 mið þy VP VP	
oð	6 + 2	6	4 + 1
	prep NP 6	oð to NP 5	prep NP 4
	oð to NP 2	oð on NP	feorra oð NP = as far as
oþ	13 + 4 + 5		
	prep NP 13		
	oþ to NP 4		
	ne VP oþ þæt VP (Mt1.25, 12.20)		
	VP oþ þæt VP (Mt14.22, 18.30, 26.36)		
oðð			4 + 11 + 4
			prep NP 4
			oðð to NP 5

			Mk 14.34 oðð vel við to NP
			oðð on NP 4
			oðð in NP
			Mk6.10 VP (við vel) oðð ðæt VP
			L12.50 VP (við vel) oðð ða hwil VP
			Lk12.59 ne VP oðð VP
			Lk20.43 [VP] oðð þætte VP
oþþ	(2)		
	(conj 2)		
oþþæ	(1)		
	(conj)		
oþþe	1 + (42) + 5		
	(conj 42)		
	14.36, 18.21 Adv = even		
	27.45 prep NP		
	1.17 oþþe to NP 2		
	13.30 oþþe to NP		
eþþa	(2)		
	(conj 2)		
oþðe	(2)	(1)	
	(conj 2)	(conj)	
oðþe	(1)		
	(conj)		
oððe		2 + (2) + 5 + 1	(5)
		(conj 2)	(conj 5)
		prep NP 2	
		oððe on NP 5	

		L12.50 oððe ða hwyle VP	
oþþa	(4)		
	(conj 4)		
oðða		1	(2)
		Mk14.32 VP oðða ða hwile VP	(conj 2)
oððæ		1	(2)
		Lk12.59 ne VP oððæ VP	(conj 2)
oððæt		1+17	11
		ne VP oððæt VP (Mk 9.1, 13.30, L9.27, 13.35, 21.32, 22.16, 22.18, 22.34)	VP oððæt VP (Mt22.44, Mk12.36, L13.21, 15.4, 15.8, 21.24)
		Mk12.36 VP oððæt VP (Mk12.36, L13.21, 15.4, 15.8, 19.13, 21.24, J21.22, 21.23)	Mt26.36 VP oððæt VP = while
		L2.15 oððæt in NP	ne VP oððæt VP (Mk9.1, 13.30, L13.35)
		L20.43 [VP] oððæt VP	Mk14.32 VP oððæt vel ða huil VP
oþþæt	5		
	VP oþþæt VP (Mt 2.13, 10.11, 13.33, 22.44)		
	Mt5.18 oþþæt VP, VP		
oþþætte	1		
	18.34 VP oþþætte VP		
oþþætti	1		
	2.9 VP oþþætti VP		
oðþæt			4
			VP oðþæt VP (L 19.13, J21.22, 21.23)

			L22.18 ne VP oðþæt VP
oðþætte			1
			L22.34 ne VP oðþætte VP
oððþæt			1
			L21.32 ne VP oððþæt VP
oððþætte			1
			L22.16 ne VP oððþætte VP
oððdæt			1
			L9.27 ne VP oððdæt VP
wīþ	4		
	prep NP 4		
wīðe			1
			Mt34.18 VP ða huile wīðe VP
wīððy			1
			Lk24.49 VP wīððy VP = until
wið	30	23+2	51 + 24 + 20
	prep NP 30	prep NP 23	prep NP 51
			(incl. 1 each vel oð, oðð, in, betiuh, 2 from)
		Mk6.10 VP wið vel ðæt VP	
		J13.38 VP wið hwile ða VP	Adv 24 (20 wið Prep NP)
			Mt10.11, 10.23, 12.20, 16.28, 17.9, 18.30, 23.39, 24.39) VP wið VP (all for donec)
			Mt1.25, 13.3 VP wið l ða huile VP
			Mt5.26 VP ðona wið VP
			Mt14.22, 24.34 VP ða huile vel wið VP

			Mt2.9 VP wið þæt miððy VP
			Mt2.13 VP wið ðon miððy VP
			2xMt5.18 wið ða hwile VP wið ða huile VP
			Mk6.10 VP wið vel oðð ðæt VP
			L12.50 VP wið vel oðð ða hwil VP
			J13.38 VP wið vel ða huile VP
uið			6 + 2
			Adv uið to NP 2
			prep NP 6

