

**Passives and the Loss of Verb Second:
A Study of Syntactic and Information-
Structural Factors**

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
<http://www.lotschool.nl>

Cover illustration: York Minster Library. Edited photograph by Gea Dreschler.

ISBN: 978-94-6093-184-0
NUR 616

Copyright © 2015: Gea Dreschler. All rights reserved.

Passives and the Loss of Verb Second: A Study of Syntactic and Information- Structural Factors

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. Th.L.M. Engelen,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 7 september 2015
om 14.30 uur precies

door

Gretha Anthonia Dreschler

geboren op 11 september 1986

te Ede

Promotor: Prof. dr. A.M.C. van Kemenade

Copromotor: Prof. dr. B.L.J. Los (University of Edinburgh, Verenigd Koninkrijk)

Manuscriptcommissie: Prof. dr. H. de Hoop (voorzitter)
Prof. dr. O.C.M. Fischer (Universiteit van Amsterdam)
Prof. dr. M. Hundt (Universität Zürich, Zwitserland)

Acknowledgements

I gratefully acknowledge the Netherlands Organization for Scientific Research (NWO) for funding this project (project no. 360-70-370/372).

The writing of this thesis is a tale of two cities (and universities), with two corresponding sets of people that I want to thank. The setting for the first couple of years, those that I spent as a fulltime *promovendus*, was the Radboud University Nijmegen. I would like to thank my promotor Ans and daily supervisor Bettelou, without whom this thesis would never have existed. Thanks for giving me the chance, as a very young near-graduate, to embark on this journey back in 2008. Throughout the years, your expertise in the field and knowledge of the literature have been of enormous value. Thanks also for all the detailed comments on any work that I sent you and our extensive discussions of the material. You were very committed to the project, while at the same time you gave me the freedom to find my own way in the material. I'm also particularly thankful for your patience in the last couple of years when the writing was not going as fast as any of us would like because I had to combine writing with my teaching position in Amsterdam.

Thank you, Erwin and Rosanne, office mates and project mates. I was very lucky to have you as company on the journey, as we could share our experiences throughout all stages. It made doing a PhD a much less lonely affair and much more fun. Thanks also to the other PhDs at Nijmegen. My fellow PhDs at the Department – Nynke, Astrid, Sanne, Meta – thanks for your support and company on our floor of the Erasmusbuilding. Thanks to the fellow PhD students who joined me in the making of *Proefdruk/In Press*, and to the PhD students that I met over the years, at PON events, courses, and summer or winter schools. The PhD community may not be as tightly-knit as the student community but it is still very valuable.

Another important place for me on the Nijmegen campus was the *Academisch Schrijfcentrum Nijmegen*. I want to thank Joy and Inge and the team of tutors for some very enjoyable years. I learned a lot about writing and the writing process, which was not just helpful for my own writing, but would also prove to be the beginning of a new direction in my career. Joy, thank you for your support and encouragement, not just as manager of the ASN but also in other settings, like the writing support groups or PhD workshops.

I would also like to thank Susan Pintzuk, Ann Taylor, and Anthony Warner of the University of York, where I spent a couple of months as a visiting scholar. Being in York brought my research to life—as I suddenly saw the history of England and English all around me—and I'm grateful that I had the opportunity to focus on my research in this setting. Thank you for the discussions on historical syntax and information structure, and especially the feedback on my own project; it really helped me to develop my ideas at this stage of my PhD. The same thanks extend to all fellow historical linguists that I've met at conferences, especially at SHES.

The setting for the last couple of years of writing my thesis has been VU University Amsterdam. I would like to thank everyone at the Department of Language, Literature and Communication, but especially the members of the chair group of English linguistics, for their support. From the very first day I've felt at home at the VU and even though combining teaching and writing was a challenge, I am thankful for all the opportunities I've had to develop my teaching, research and organizing skills. At the VU I also found another set of PhD students and people who had just finished who I could rely on for support and company: Gudrun, Tessa, Kasper, Vaclav, Olga, Rebecca, thanks for the many chats during coffee, tea and lunch breaks, and for fresh perspectives from different branches of linguistics.

Mike, thank you for your support, not in the least for editing my manuscript and being, in effect, the first person to read the entire book (besides me and my supervisors) and giving me the confidence that it was ready to go out into the world. Margreet, thanks for your valuable advice on many occasions—not just about anything related to my thesis, but also about many teaching-related issues. And for practically pressing 'send' for me at a key moment in finishing.

Adding one further city to the list, I would like to thank the linguistics staff at the Department of English Language and Culture at Leiden University—Martina Noteboom, Colin Ewen, Marion Elenbaas, Bert Botma—for introducing me to the wonderful world of vowel diagrams, phonetic transcriptions, syntax (of course), and historical linguistics.

Finally, I want to thank everyone who kept me sane during these sometimes crazy years. These are the people who transcend any specific setting because they were there for me all the time. First of all, my family. My sisters and brothers (in law), nephews and nieces reminded me, if ever I needed reminding, that there is more to life than linguistics. Mum and Dad, thank you for your unconditional support throughout everything. Mum, you are not just the best mum ever, but you've also been the best writing coach, allowing me to check in with you to discuss (lack of) progress, and listening to my explanations of what a particular section or chapter was about when I was stuck (even though you are allergic to the linguistic terms). To my friends I'm thankful for their support, encouragement or distractions—whatever was needed—and for coffees and chats, for beers (or ciders), film/TV nights, joint working sessions, holidays, and for believing that I could finish my thesis even when I wasn't convinced of it. I am also very grateful for the BBC's daytime television that allowed me to dream of life in the English countryside when the writing on working-from-home days was tough; for Dutch radio stations Radio 1 and 3FM for adding a rhythm to my day; and for the Bagels & Beans for providing good coffee and a nice *and* productive escape from the home office when necessary. And finally, I would like to thank all the people I've met over the years—at birthday parties, at the gym, in shops or anywhere—who were interested in the topic of my project or just the fact that I was doing a PhD and who made me realize how lucky I was to be working on something that I loved.

Contents

1. Introduction	13
Section 1.1 Background	14
1.1.1 The exceptional nature of the passive in English	14
1.1.2 Verb second and its information-structural aspects	16
1.1.3 Consequences of the loss of verb second: proposals by Los and Seoane ..	18
1.1.4 The problem of interaction between syntax and information structure from a Minimalist perspective	21
Section 1.2 Research question and plan	23
Section 1.3 A note on method and methodology	25
Section 1.4 Outline	27
2. The Information-Rearranging Potential of Passives and Verb Second ..	31
Section 2.1 Introduction	31
Section 2.2 Old English verb second and its loss: information-structural aspects	32
2.2.1 Categorical and variable verb second in Old English	33
2.2.2 A time frame for the loss of verb second	39
2.2.3 The pragmatic consequences of the loss of verb second	43
2.2.4 Summary and conclusions	48
Section 2.3 Syntactic and functional properties of the English passive	49
2.3.1 A definition of the passive	49
2.3.2 The information-rearranging function of the passive	57
2.3.3 A Minimalist analysis of the passive	61
2.3.4 Summary and conclusions	68
Section 2.4 Operationalizing information-rearranging functions	69
2.4.1 Given-before-new as the unmarked order of information	69
2.4.2 Given-before-new as a function of syntactic variation	77
2.4.3 Annotating givenness in a corpus	83
2.4.3.1 Annotation schemes for PDE	83
2.4.3.2 Annotation schemes for older (Germanic) languages	87
2.4.4 Summary and conclusions	92

3.4.3.1	The ECM active and passive with <i>believe</i> verbs.....	168
3.4.3.2	Verbs of commanding and permitting.....	170
3.4.3.3	Related passives	174
3.4.3.4	Conclusion	177
3.4.4	Summary and conclusions	178
Section 3.5	Conclusion	179
4.	<i>The Given-before-New Principle in Verb-Second Old English</i>	183
Section 4.1	Introduction	183
Section 4.2	The passive in Old English	185
4.2.1	A characterization of the Old English passive	186
4.2.2	A corpus study of passives in YCOE	193
4.2.2.1	Method of selection and analysis	193
4.2.2.2	The order of information in long passives	198
4.2.2.3	Passive subjects as discourse linkers	204
4.2.2.4	Late subjects.....	211
4.2.3	Summary and conclusions	216
Section 4.3	Object fronting	217
4.3.1	Method and overall frequencies	219
4.3.1.1	Other object-subject orders	221
4.3.1.2	Frequency of object fronting in the database	224
4.3.1.3	Information status and anaphoricity categories, including ‘speech’, ‘event’ and ‘topic’	225
4.3.2	Relative information status of fronted object and subject	227
4.3.3	The first position as a local anchoring position	233
4.3.4	Inversion in object fronting sentences	236
4.3.5	Summary and conclusions	241
Section 4.4	Clause-initial PPs.....	242
4.4.1	Method and selection	243
4.4.2	Clause-initial PPs as local anchors	247
4.4.3	Inversion in PP-initial clauses	253
4.4.4	Information-structural factors in <i>Orosius</i>	259
4.4.5	Summary and conclusions	264
Section 4.5	Conclusion	265
5.	<i>Aspects of the Loss of Verb Second</i>	267
Section 5.1	Introduction	267
Section 5.2	The loss of inversion in PP-initial clauses	269

5.2.1	Locative inversion in PDE: effect of PP and verb _____	270
5.2.2	Inversion after clause-initial PPs in Early Modern English _____	275
5.2.2.1	Method and selection.....	275
5.2.2.2	The loss of inversion after PPs.....	282
5.2.2.3	Information-structural factors in subject placement in E1.....	289
5.2.2.4	Conclusion	294
5.2.3	Remnant inversion patterns in Modern British English _____	295
5.2.4	Summary and conclusions _____	299
Section 5.3	The loss of local anchoring	300
5.3.1	The increasingly subject-initial nature of English _____	301
5.3.2	Clause-initial PPs: a decrease in local anchoring and an increase in contrastive and frame-setting uses _____	307
5.3.2.1	Loss of <i>P+D/PRO</i> in clause-initial position	307
5.3.2.2	A study of the information-structural character of the clause-initial PPs in E1	311
5.3.2.3	Conclusion	317
5.3.3	The increasing restrictions on object fronting _____	317
5.3.4	Summary and conclusions _____	323
Section 5.4	Conclusion	323
6. Responses to the Loss of Verb Second.....		325
Section 6.1	Introduction	325
Section 6.2	The development of the passive after ME	326
6.2.1	The frequency of short and long passives _____	328
6.2.1.1	Frequency of the passive.....	328
6.2.1.2	Frequency of long passives.....	336
6.2.1.3	Conclusion	339
6.2.2	The relevance of the passive as an information re-arranger after ME ____	339
6.2.2.1	The subject-creating or discourse-linking function of the passive.....	340
6.2.2.2	The passive as a topic introducer	344
6.2.2.3	Conclusion	349
6.2.3	Summary and conclusions _____	349
Section 6.3	Development of the new passives after ME.....	350
6.3.1	Prepositional passive _____	351
6.3.2	Recipient passive _____	356
6.3.3	ECM passive _____	363
6.3.4	Summary and conclusions _____	371
Section 6.4	Middles and non-agentive subjects	371
6.4.1	Middles _____	373

6.4.1.1	Defining middles.....	373
6.4.1.2	New middles?	377
6.4.1.3	Middles diachronically.....	379
6.4.1.4	Conclusion	387
6.4.2	Non-agentive subjects	387
6.4.3	Summary and Conclusions	392
Section 6.5	Conclusion	392
7. Conclusion.....		395
Section 7.1	Summary of findings per chapter	396
Section 7.2	Answers to the main questions	400
7.2.1	The interaction between syntactic and information-structural factors in Old English	400
7.2.2	Changes in the interaction between syntax and information structure after the loss of V2.....	401
7.2.3	Responses to the loss of V2	402
7.2.4	Conclusion: a complex interplay of syntactic and information-structural factors	403
Section 7.3	Suggestions for further research	405
Bibliography		409
Samenvatting (Summary in Dutch).....		423
Curriculum Vitae		432

1. Introduction

This thesis investigates the interaction between syntax and information structure within a Minimalist model of grammar. As an aspect of language use, information structure has not traditionally been a focus of investigation within generative syntax, but recent years have seen an increased interest in exactly this aspect of language and its influence on word order. In this thesis, I aim to shed light on the question of interaction between syntactic and information-structural factors by investigating two developments in the history of English that have been shown to be influenced by both syntax and information structure: verb second (V2) and the passive. The loss of V2 in the fifteenth century—a dramatic change in the syntax of English with far-reaching consequences for the possibilities left to reorder arguments, and hence information, in the sentence—incorporates both syntactic and information-structural changes. At the same time that V2 was lost, significant changes also took place in the passive: not only was there a remarkable increase in types of passives around 1400, but the use of passives overall increased from the sixteenth century onward. One of the central properties of the passive is that it reorders arguments, which leads to changes in the order of information in a sentence—a function which the V2 system in Old English (OE) also performed. This raises the question whether the two developments are related through their syntactic and information-structural properties. In this thesis, I investigate the relation between the syntactic and information-structural factors that play a role in the loss of V2 on the one hand and the changes in the passive on the other hand, through four case studies, spread across the periods of the history of English, from the ninth century until the present day. The main conclusion that I will draw is that there is clear evidence of the influence of syntax on information structure, but only limited evidence of influence

from information structure on syntax. Although information structure has a major influence on the choice between options that the syntax makes available and the frequency and use of particular positions and constructions, it does not influence the syntactic options that are available, in that it does not trigger syntactic innovations.

Section 1.1 Background

The proposed connection between the passive and V2 in the history of English as outlined above builds on their shared information-structural function: they both allow for a reordering of arguments in the sentence to achieve a particular effect and both go through quite a remarkable development, as can be seen in the increase in types of passives and the loss of V2. At the same time, the constructions are syntactically fundamentally different. The syntactic developments, when combined with the shared information-structural function, lead to a scenario in which the loss of one construction—V2—affects the use of another construction—the passive.

1.1.1 The exceptional nature of the passive in English

The passive in Present-day English (PDE) is exceptional in two respects: first, it has been observed that the passive is more frequently used than in closely-related languages (e.g. Siewierska 1984: 218, Halliday 2000); second, English has a wider range of passive constructions than its West-Germanic sister languages, including types that are rare cross-linguistically, as illustrated in (2)-(4), in addition to the standard passive in (1).

- (1) The book was written by John.
- (2) Mary was given the book.
- (3) The doctor was sent for.
- (4) John was believed to be a liar.

In the regular, direct passive in (1), the direct object of a transitive verb is passivized and becomes the subject of the passive sentence. In (2) the indirect object of a ditransitive verb has been passivized, resulting in what I will call, following Allen (1995), a recipient passive. Example (3) shows a prepositional passive, in which the complement of a preposition (*for the doctor*) has been passivized. Finally, the example in (4) illustrates the ECM passive, a term based on that for its active counterpart (*They believed John to be a liar*) which has become known as the Exceptional Case-Marking construction in Government and Binding theory (see Chomsky 1981).

The passives in (2)-(4) were not present in all earlier stages of the language. In fact, until roughly 1400 AD, English had only three types of passives: the direct passive as in (1); the impersonal passive as in (5); and a marginal morphological

passive as in (6).

- (5) *ac him næs getiðod ðære lytlan lisse*
 but him.DAT not-was granted that small favour.GEN
 ‘But he was not granted that small favour’
 (ÆCHom I 23.330.29; Denison 1993: 108, his 29)
- (6) *and se munuc hatte abbo*
 and that monk was-called Abbo
 (ÆLS II 32.3; Denison 1993: 421, his 33)

It is only around 1400 that the new passives in (2)-(4) begin to appear in the language, a development which has been documented in detail in various studies—for instance in the voluminous works on OE syntax by Visser (1963-1973) and Mitchell (1985). The following examples show the first attested examples, including their date, of the prepositional passive (7), recipient passive (8), and ECM passive (9), based on a detailed (re)examination of the data by Denison (1985, 1993), Allen (1995) and Warner (1982).

- (7) *þer wes sorhe te seon hire leoflich lich faren so reowliche wið*
 there was sorrow to see her dear body dealt so cruelly with
 (c1225; *St. Juliana* (Royal) 22.195; Denison 1993:125, his 10)
- (8) *Item as for the Parke she is a lowyd Every yere a dere*
 ‘Item: as for the park, she is allowed a deer each year’
 (1375; *wardBlount* p.205; Allen 1995: 393, her 256)
- (9) *þat man is seid to have a fend whom þe fend disseyveþ, as he is seid to have an heed þat is heded bi þis hede;...*
 ‘...as the man who is dominated by a head is said to have this head.’
 (c1400; *Serm* i.125.2; Warner 1982: 136, his 9)

The prepositional passive is the first of the new cross-linguistically rare passives to occur, with examples dating to the early thirteenth century, while the first examples of the recipient and ECM passive are from the late fourteenth century. The number of examples of prepositional passives in the thirteenth century, however, is limited, which means that these three new and rare passives are introduced within the span of roughly one century. After their introduction, all three passives gain foot in the language, becoming more frequent at different rates over the following centuries, and they remain not only grammatical but also productive in the English language up to the present day (see the studies cited above, and Seoane 1999b, Los 2005, 2009).

It is not just in types of passivization that the English passive shows a

remarkable development, but also in terms of frequency. Seoane (1999a, 2000, 2006) has shown that there is an increase in the use of passives compared to active clauses in the fifteenth and sixteenth century. She finds that the percentage of passives with respect to active transitive clauses—i.e. those transitive clauses which can be passivized—increases from 18.0% in M4 (1420-1500) to 24.1% in E3 (1640-1710) (Seoane 2006: 371, Table 15.1). In addition, several authors have stated that PDE uses the passive to a greater extent than its closely related languages (Siewierska 1984, Halliday 1967, 2000, Hawkins 1986), although I have found no publications which provide evidence for this from corpus studies. This prominence of the passive in PDE, and the increase in frequency in the fifteenth and sixteenth century can be seen as another indication for what could informally be called ‘the rise of the passive’.

1.1.2 Verb second and its information-structural aspects

The major changes in the passive take place during the same period as one of the major syntactic changes in English since its recorded history: the loss of V2. The V2 constraint, a characteristic of modern Germanic languages such as German and Dutch (see Vikner 1995, Zwart 1997), entails that the finite verb in main clauses always occurs in the second position in the clause, leading to inversion of subject and finite verb in those sentences that have a non-subject in clause-initial position, as in (10).

- (10) *þa gemette he ðær ænne þearfan nacodne*
 then met he there a beggar naked
 ‘Then he met a poor man, naked’
 (ÆLS (Martin) 61—62; Los 2009: 100, her 2)

This sentence starts with the adverb *þa*, which is followed by the finite verb *gemette*, in turn followed by the subject *he*. The standard analysis for modern V2 languages is that the first element occurs in Spec,CP and the finite verb in C (Vikner 1995) in all main clauses. Matters in OE are not nearly as straightforward, however. In the literature, most authors now make a distinction between two types of V2 (see e.g. Haeberli 2000, Fischer et al. 2000). The first type—illustrated in example (10)—is similar to German and Dutch, in that inversion invariably occurs: in clauses beginning with an operator, a negator or the discourse adverbs *þa* and *þonne*, the finite verb always occurs in second position. This type of inversion survives into PDE as question inversion (*Where **did you** buy that dress?*) and subject-auxiliary inversion (*Never before **have I** been this happy*).

The second type of V2 is more variable and occurs in sentences which have a non-operator in clause-initial position. The following examples illustrate how, when the first element is a prepositional phrase (PP) (or indeed, an object or adverb), subjects can occur either before or after the finite verb.

- (11) *Mid þusend searocræftum wolde se swicola deofol þone halgan wer on sume wisan beswican,*
 ‘With a thousand wily arts did the treacherous devil strive in some way to deceive the holy man,’
 (o3, coelive,ÆLS_[Martin]:710.6427)¹
- (12) *ON sumere tide Martinus stah to anre upflora,*
 ‘On one occasion Martin was mounting to an upper floor,’
 (o3, coelive,ÆLS_[Martin]:601.6358)

The usual analysis for these sentences is that the verb does not move to the highest head in the clause (C in CP), but to a lower position. In split-IP proposals this position is identified as AgrSP (Haeberli 2000, Warner 2007), or FP (Fischer et al. 2000), a functional projection on top of TP. This in turn leads to a situation where there are two landing sites for subjects in the IP—either before the finite verb in FP (Spec,FP) or after it (Spec,TP)—giving rise to surface verb-second or verb-third sentences. Note that even though the order of PP, subject, and verb in (12) is similar to that in the PDE translation, the syntactic analysis is different: the PDE verb is in T while the OE verb is in AgrS/F, a position which is lost when V2 is lost.

One of the factors that determine whether the subject occurs before or after the finite verb in this second type of V2 is whether it is pronominal or nominal: pronouns always occur before the finite verb, while nouns generally follow it (Van Kemenade 1987, Pintzuk 1991, 1993). While several authors have tried to find other syntactic factors that can explain the variation in subject positions, none of their accounts are able to explain all variation. Recent proposals, such as Bech (2001), Hinterhölzl & Petrova (2010) and Van Kemenade & Westergaard (2012), have proposed an information-structural explanation for the word order variation in the OE V2 system: they propose that discourse-old subjects occur before the finite verb, while discourse-new subjects occur after the finite verb.

Another information-structural aspect of the V2 organization of main clauses in OE is that it entails a different function for the clause-initial position. Specifically, it has been proposed that the clause-initial position is mainly used to encode discourse-old information and to provide a link to the immediately preceding discourse, a function which has been called “local anchoring” (Los 2012, Los & Dreschler 2012; for similar proposals, see Van Kemenade & Westergaard 2012 and Hinterhölzl & Petrova 2010). This function is illustrated in (13).

- (13) ‘They would then have soon perished before their enemies, if they had not broken into the city by a device, which was most shameful, though it was afterwards thought most worthy of them; [...]

¹ Examples taken directly from the YCOE are given according to their file name (all starting with “co...”). The superperiod is included before the file name (see Section 1.3).

Pysne nyttan cræft, þeh he arlic nære, funde heora tictator, Camillis hatte.
 this useful device, though it honourable not-was, found-out their dictator,
 Camillus called
 ‘This useful device, though it was not honourable, was found out by their
 Dictator, Camillus.’
 (coorosiu,Or_2:8.51.24.983)

What makes the clause-initial position in OE stand out is that it is used to express these local anchors as well as marked, or contrastive links, a function it retains in PDE. These marked and unmarked information-structural functions can be expressed through a range of syntactic means, including both adjuncts—such as the PPs in (11) and (12)—and arguments, such as the contrastive objects in (14).

- (14) ‘Then went he through all the Egyptian land, sowing God’s seed, and healed the sick.

Hreoflige he geclænsode fram ðære unclænan coðe, Wode he gehælde
 lepers he cleansed from the unclean disease, possessed he healed
 ‘He cleansed the lepers from the unclean disease; he healed the possessed’
 (o3, coelive [Mark]:6.3189)

While the marked, contrastive use of the initial position survives into PDE, the unmarked discourse-linking function which is characteristic of OE V2, and indeed, modern V2 languages, is no longer possible in PDE: PPs and objects in first position are overwhelmingly contrastive, or used as frame setters, elements which provide the context for a following statement, as in (15) (see Los 2009, 2012; Hasselgård 2010; Krifka 2007; Speyer 2010).

- (15) A: How is John?
 B: **Healthwise / As for his health**, he is fine.

(Krifka 2007, his 47)

Fundamentally different, syntactically, from PDE SVO, OE builds on a different interaction between word order and information-structural principles. The question is when the use of the first position as a ‘local anchoring’ position changes and what happens to the information-structural aspects both of subject placement and of the first position when V2 is lost.

1.1.3 Consequences of the loss of verb second: proposals by Los and Seoane

The recent approach to the analysis of V2 in terms of information structure has also led to a new view on the loss of V2 and specifically on the consequences for the language when the system was lost. Los (2005, 2009, 2012) provides a scenario in which the loss of V2—which for the major part took place in the fifteenth century—

limited information structure in such a way that new syntactic constructions arose as a result, specifically the cross-linguistically rare recipient, prepositional and ECM passives. Seoane (2006) follows with a proposal which also connects the passive to the loss of V2: she ascribes an increase in the use of passives with respect to actives in Early Modern English (EModE) to this same scenario of information-structural consequences of the loss of V2. Specifically, both Los and Seoane propose that with the loss of V2, the subject becomes the only way to begin a sentence in an unmarked way, thereby creating the need for more strategies to place an argument in subject position.

Los's proposal builds on the information-structural character of the first constituent in OE as a V2 language, based on, among other things, observations about the first position in modern V2 languages like German and Swedish (she refers to Bohnacker & Rosén 2007). Specifically, she proposes that OE had two information-structurally unmarked positions for discourse-old information at the beginning of the clause: the presubject position and the subject position. Both these positions can host old information and by virtue thereof link the sentence to the previous discourse, as illustrated in the following example.

- (16) *Be þære he gestrynde ðry suna Her & Onam & Sela.*
 By that he begot three sons Er & Onan & Shelah.
 'He had three sons with her: Er, Onan and Shelah.'
 (cogenesiC,Gen_[Ker]: 38.3.83; Dreschler & Hebing 2011: 59, their 5)

In (16), both the clause-initial PP *Be þære* and the subject *he* represent discourse-old information, and thus provide a link to the preceding discourse.

The situation in PDE is decidedly different: Los (2012) and Los & Dreschler (2012) present suggestive evidence that the clause-initial position has become "pragmatically marked" and that the subject position is the only remaining position for unmarked themes. For instance, clause-initial PPs containing typical discourse-old elements such as pronouns show a decline (Los & Dreschler 2012: 866). The consequence of this development is that the subject becomes increasingly important as a clause-initial element. It also takes over part of the 'discourse-linking' function that the first constituent could before perform, under the assumption that discourse-linking and expression of unmarked themes remain important functions in the language (see Komen et al. 2014).

Los (2009) presents two different types of consequences of the increased importance of the subject which will be central to this thesis. On the one hand, she points to passives, which she describes as the "prototypical way" to place an argument in subject position. The new passives were introduced around the time that V2 was lost, and the use of these passives and of passives in general increased in Early Modern English, the period immediately following the loss of V2, as reported by Seoane (2000, 2006). Los's original proposal is built around the ECM passives, whose motivation, she claims, is "overwhelmingly information-structural" in PDE

(2009: 119). Consider the following examples provided by Los, which would all be a grammatically possible follow-up to a sentence in which several names are introduced; the old information is given in bold.

- (17) **Many of them** are generally assumed to be Americans.
- (18) People generally assume that **many of them** are Americans.
- (19) It is generally assumed that **many of them** are Americans.
- (20) **Of these people** it is generally assumed that they are Americans.

(Los 2009: 120, her 36-38)

While all options are grammatically possible, it is example (17) that provides the best information-structural transition, in encoding the discourse-old *many of them* as the clause-initial subject. In (18) and (19), the discourse-old information is too far from the initial position; and in (20), it is encoded as a marked initial PP. In contrast, this last example would be a good information-structural fit in OE, where the initial PP would be unmarked. In a later publication, Los (2012) also includes the prepositional passive and recipient passive in this same scenario. Another type of ‘subject strategy’ that is described by Los (2009) as a consequence of the increased importance of the subject is the range of non-agentive subjects that is used in PDE, as discussed in detail by Rohdenburg (1974). Example (21) illustrates.

- (21) 2004 saw the advent of routers with turbo modes.

(Los 2009: 118, her 33a)

The subject in (21), the year *2004*, is not the agent of the verb *saw*. Nevertheless, this non-agentive subject is allowed in PDE, which has the effect that this argument now occurs in subject position, the position for unmarked themes, instead of being expressed as an adverbial *In 2004*. Los proposes that this range of strategies that can place an argument in subject position is a result of the loss of V2, or more specifically, the increased importance of the subject and increased marked nature of non-subjects in initial position.

Seoane (2006) also proposes a relation between the loss of V2 and the passive, but in a somewhat different way, focusing not only on the subject-creating function of the passive, but also on the information-rearranging function of long passives. In earlier work, Seoane (2000) reports an increase in the use of passives in EModE and shows the importance of the (long) passive in reordering the arguments in the sentence, thereby ‘restoring’ the neutral given-before-new order of information in the clause. Consider the context in (22) and the following active and passive sentences in (23)a and b respectively.

- (22) John crossed the street

- (23) a. A car hit him.

- b. He was hit by a car.

In the active sentence, the information flow is disrupted because a new referent, *a car*, precedes the old referent *him*. This flow is restored in the passive sentence, which allows *He* to be the subject and occur in the unmarked position for old information in the clause. In a later publication, Seoane (2006) includes Los's scenario in her study of this information-rearranging function of the passive. Where Los focuses on passives as subject strategies, Seoane focuses on the rearranging function of the long passive and achieving a given-before-new order of information. She describes the passive as "a word order rearranging strategy which can, in addition, create subjects/unmarked topics" (370), whose relevance increased because of the new importance of subjects as unmarked themes after the loss of V2.

Two scenarios present themselves on the basis of these two proposals and the other data we have about the loss of V2: (i) a causal relation between the loss of V2 and the introduction of new passives and new subject strategies; (ii) a looser relation between passives and the loss of V2, in that what is important are the consequences of the loss of V2, rather than the loss of V2 itself, since those consequences create the need for subject-creating and information-rearranging strategies. The two proposals make different predictions for a syntactic model. In the first scenario, which is the one that Los proposes, an information-structural need leads to a syntactic change: "syntactic change (the loss of verb-second) affects discourse, and pressure from discourse results in syntactic change (i.e. the emergence of new constructions)" (2009: 121). The predictions from the second scenario—corresponding to Seoane's observations—imply that the syntax stands on its own and the differences over time only affect the frequency of use of a construction. This latter scenario can gain even more strength if we take into account the fact that there are constructions that were lost in English that are likely alternatives to the passive for expressing the information-rearranging function, such as the impersonalization strategy *man* (see Los 2002) and rearranging strategies of object fronting (see Speyer 2010) and to a lesser extent PP-fronting.

1.1.4 The problem of interaction between syntax and information structure from a Minimalist perspective

Within a Minimalist model of grammar, it is not at all evident that syntax and information structure interact. Language use has not played a major role in the models of grammar in the generative syntactic tradition, sparked off by Chomsky's (1957) *Syntactic Structures*. Chomsky made famous the distinction between competence and performance, and states explicitly in *The Minimalist Program*—the book that was the start of this latest version of generative syntax—that the aim of the program is to build a model of competence in language; it is not to account for instances of performance. In fact, Chomsky speaks of describing or explaining language in "the idealized conditions of a homogeneous speech community" (1995: 19). Moreover, he proposes focusing on the core system ("pure instantiations of U[niversal] G[rammar]"), leaving more peripheral matters like "historical accident, dialect mixture, personal idiosyncrasies, and the like" aside (*ibid.* 19-20). Generally

speaking, issues of language use belong more to the periphery than the core, more to performance than competence.

Despite these aims, factors of language use or discourse have in fact been included in several Minimalist analyses, for the simple reason that in some cases the influence of discourse elements cannot be ignored as a determining factor in the use of a construction or in explaining variation. Two examples of constructions where discourse notions have featured are Scandinavian Object Shift and Scrambling (Richards 2004: 4-5, see also Chomsky 2001): in both Object Shift and Scrambling, there is variation in the position of the object; crucially, when it is moved leftward, it receives a defocalized (or old/topical) reading. Richards (2004) argues that in these cases, there is an optional EPP-feature, which is motivated by these interpretative effects and in turn licenses the movement. In other words, a different selection of elements in the numeration—in this case the selection of an EPP-feature—leads to interpretive effects, although the feature itself is not defined in terms of these effects. In a different type of solution, Aboh (2010) has proposed that information-structural notions should be translated as features in the derivation, just like case or person. He illustrates this with the following examples:

- (24) I like John very much.
- (25) John, I like him very much.
- (26) John I like very much.

(Aboh 2010: 15, his 4, 3a and 3c)

While (24) is the basic order and John receives a neutral reading, in the left-dislocation in (25), *John* receives a topic reading and in the focus movement sentences in (26), *John* receives a focus reading. Aboh states that these readings are “additional meanings” which the sentences acquire, which cannot be “directly ascribed to syntax” (2010: 17) and that syntax cannot be responsible for the observed discourse effects. Aboh proposes, on the basis of data from Gungbe, that Topic, Focus and Interrogative Force should have their own features (and own projections), because Gungbe has lexical elements (i.e. syntactic positions) whose only function is to mark topic or focus. Following the principles of UG, he takes this as evidence that these features are present in other languages as well.

In addition, following Rizzi's (1997) influential work, syntactic approaches have included reference to information structuring in the representation of clauses, i.e. not just acknowledging the relevance of information-structural factors but specifically including it in the syntax. In this so-called ‘cartographic’ approach, the CP-level of the clause is divided into several functional projections, the contents of which correspond to specific information-structural notions, most notably topic and focus. Abstracting away from the details of the features and positions involved, movement to these positions takes place for information-structural purposes on the basis of a stipulated formal feature; for instance, the topic of the clause moves to the specifier position of the Top[*ic*]P. This approach includes the conviction that

information structure determines word order, and also seems to assume that at least parts of it—notions like topic and focus—are restricted enough to describe in sufficient detail, i.e. that there are specific aspects that can be shown to be the driving force for movement. The attraction is that information structure can explain word order variation and processes in the language that have previously not received a satisfactory explanation, such as topicalization or left-dislocation. This has led to a large body of literature on the interaction (or mapping) between syntax and information structuring, and a wealth of information-structurally motivated projections in especially the Left Periphery (for OE, see Hinterhölzl & Petrova 2010 and Walkden 2015).

In short, the connection between syntax and information structure presents a challenge in a Minimalist framework, although several views on acknowledging information structure as an influence, or including it in the derivation, have been proposed. The necessary first step, however, in any undertaking to assess their interaction based on a particular case study—as indeed many studies have already aimed to do—is to establish how important information-structural factors are in explaining particular instances of grammatical variation and change, or the use of a particular construction.

Section 1.2 Research question and plan

Based on the developments and questions outlined in the previous section, the main aim of this thesis is (i) to investigate the relation between the loss of V2 and the ‘rise’ of the passive and other subject strategies in the history of English and (ii) to identify the syntactic and information-structural factors involved, with the aim of determining to what extent syntax and information structure interact. In order to answer these questions, the body of this thesis consists of four case studies which investigate the information-structural and syntactic factors of different constructions at various stages of the language. The case studies combine analyses of syntactic patterns and structures with analyses of information-structural motivations, based on the annotation of information status.

The first case study, focusing on fourteenth-century material, investigates the introduction of three new passives—prepositional, recipient and ECM—and aims to determine the time of their introduction with respect to the loss of V2. It provides the basis for the other case studies by establishing when and why these passives were introduced. The results of this study provide the data to establish whether there is a close relation between the loss of V2 and the developments of the passive, i.e. whether the introduction of these passives follows the loss of V2. In addition to determining the time of introduction with respect to V2, this case study also considers the possibility of a structural relation between the changes that make up the loss of V2 and the introduction of the new passives. Although the introduction of these passives has been described in quite some detail, the syntactically annotated corpora that have since become available (see Section 1.3) make it possible to trace

the developments in a more systematic way. These additional data can shed more light on the reasons for and the time of the introduction of these cross-linguistically rare passives.

The second, third and fourth case studies together aim to provide an account of the relation between the passive and V2 in a different way, focusing on the information-structural aspects of V2 and the information-structural consequences of the loss of V2 from a more general perspective. The presence of object and PP preposing in OE, of the type that places unmarked old information in initial position, in addition to the strategies that remain in PDE, such as subject-initial clauses and passives, points to the possibility that a similar function, that of creating a given-before-new order of information, was expressed by a wider range of constructions in OE than in PDE. This suggests both a less prominent role for the passive in OE than it has in PDE, and a scenario where the loss of these alternatives leads to an increased importance of the passive. This means a scenario where the loss of V2—to which we assume these constructions to be crucially connected—led to an increased importance of the passive. In addition, the increased importance of the subject, also an effect of the loss of the presubject first position for unmarked themes, could also have led to an increase in the use of passives and other subject-creating constructions. In order to establish the likelihood of this scenario, we need to gather evidence in the shape of the relative time frames of the developments and their relation to each other, both syntactically and functionally.

The second case study, therefore, looks in detail at the information-structural aspects of V2 as it is found in OE, and aims to determine to what extent different word orders are determined by information-structural factors. It investigates the information-rearranging properties of passives, object fronting and PP preposing—which all rearrange arguments in the clause—as well as determining the information-structural properties of the clause-initial position in this range of constructions. The information status of the subject is considered in detail because of recent proposals (Hinterhölzl & Petrova 2010, Biberauer & Van Kemenade 2012) that subject placement in OE is information-structurally determined.

The third case study specifically looks at how the two information-structural aspects of V2—rearranging possibilities and use of the first position—change after the OE period. This case study is based on data from 1500 up to the twentieth century, but focuses on the sixteenth century, the period immediately following the end date usually given for the loss of V2. It focuses on the use of the first position and the properties of the subject, again investigating both the syntax as well as the information structure. The case study aims to trace the increasingly subject-initial nature of English and aims to establish what the information-structural effect of this is on other clause-initial elements. More specifically, it investigates whether these clause-initial elements indeed became marked and if so, when this aspect of the loss of V2 took effect.

The fourth case study, finally, investigates possible consequences of the loss of V2 in the form of several subject-creating strategies that are introduced or increase

in use after the sixteenth century. First, it considers long passives as creators of given-before-new order. Second, it determines the status of several strategies as subject-creators and traces their development: (i) the new passives; (ii) middles (*Bureaucrats bribe easily*) and (iii) non-agentive subjects (*This year saw the introduction of 4G networks*). It considers the information-structural function of all these subject-creating strategies as well as their frequency in the corpora in the periods after the loss of V2.

The combined insights from the first case study on the relation between V2 and the passive on a syntactic level, and the second, third and fourth case studies on a more general information-structural level, are used to establish the interaction between syntax and information structure in these developments. Because the influence of syntax on information structure seems clear, I specifically aim to assess whether there is any evidence of information structure influencing syntax, i.e. changes in the structural options that are available to speakers of a language. In such a scenario, syntax and information structure can really be shown to interact and work together.

Section 1.3 A note on method and methodology

The corpus studies of diachronic English are based on the following syntactically annotated corpora that have become available in recent years:

- the *York-Toronto-Helsinki Parsed Corpus of Old English Prose* (YCOE; Taylor, Warner, Pintzuk & Beths 2003);
- the *Penn-Helsinki Parsed Corpus of Middle English*, second edition (PPCME2; Kroch & Taylor 2000);
- the *Penn-Helsinki Parsed Corpus of Early Modern English* (PPCEME; Kroch, Santorini & Diertani 2004);
- the *Penn Parsed Corpus of Modern British English* (PPCMBE; Kroch, Santorini & Diertani 2010).

All examples in this thesis follow the standard format for references from these corpora; detailed documentation on the texts that are included in the corpora is available on their websites.² Table 1.1 shows the periods covered by these corpora—based on the conventional division of the stages of the English language—as well as the subperiods and their abbreviations.

² For the YCOE, this is <<http://www-users.york.ac.uk/~lang22/YcoeHome1.htm>>; the other corpora can be accessed via <<http://www.ling.upenn.edu/hist-corpora>>

Table 1.1 Period divisions for the historical corpora of English

Old English (OE)		Middle English (ME)		Early Modern English (EModE)		Modern British English (MBE)	
O1	450-800	M1	1150-1250	E1	1500-1569	B1	1700-1769
O2	850-950	M2	1250-1350	E2	1570-1639	B2	1770-1839
O3	950-1050	M3	1350-1420	E3	1640-1710	B3	1840-1914
O4	1050-1150	M4	1420-1500				

For most studies, the entire corpora are used, but the PPCEME needs a bit more explanation: it consists of a first edition—the Helsinki part—and two additions (p1 and p2); for each study, it will be indicated which part of the corpus was used. The corpora were searched with the programme *CorpusStudio* (Komen 2011), which builds on the search programme *CorpusSearch* that originally came with the corpora (Randall 2005-2007). Throughout this thesis, I will specify in as much detail as possible (and relevant) how the different results were achieved by giving the queries that were used.

The use of historical corpora is not without its drawbacks, as has been widely discussed in the literature (for a discussion of the corpora used here, see e.g. Kohonen 2007, Kytö & Pahta 2012). In order to avoid some of the common problems, this study consists of a combination of larger quantitative studies and smaller-scale detailed qualitative studies. Where necessary, the findings are supported by data from other corpora. The main limitation of corpora for the present study is that some of the passives, middles and non-agentive subject constructions are either not frequent, or to a large extent lexically determined. The syntactically annotated corpora turn out to be too small to provide a reasonable number of examples for a substantial corpus study. In these cases, I have tried to build on other studies or other corpora, but as is the case for historical corpora more generally, we sometimes have to work with whatever is available and the only solution is to treat the results with care. Where possible and appropriate, statistical tests have been applied to consider the validity of the results.

An added difficulty in this study is the coding of information structure in a historical corpus, which means that it is necessary to make assumptions about the writer's and reader's knowledge, which, up to more than 1,000 years after the texts were written, is not straightforward. Moreover, the numerous annotation schemes that are available for modern languages show that even with the help of speakers, decisions on information structure are not always easy to make. Nonetheless, there is sufficient knowledge of the historical context at the time that the texts in the corpora were written to be certain of at least the general notions which most people were familiar with. In addition, in recent years several researchers and project groups have undertaken the task of annotating information structure in OE, which means there is now considerable consensus on at least the basics of such an annotation (cf. Bech 2001, Komen et al. 2014, Haug et al. 2014, Taylor & Pintzuk

2014). One of the questions in this thesis, in fact, is to what extent finer-grained distinctions play a role in the constructions—something which also has a bearing on the study of information structuring in modern languages. Due to the nature of the investigation, the numbers for the information-structural information are relatively small. I do believe, however, that they are representative enough for generalizations to be made. Moreover, the focus on a smaller selection of texts allows for a detailed study of the entire text and more precise annotations than a larger-scale study could achieve.

In this thesis, I will use a minimal version of Minimalist syntax, as outlined in Chomsky (1995, 2000) and adapted in various syntactic publications since—see Biberauer & Roberts (2005) for an application to OE clause structure. This means that I will use the basic principles but at the same time will simplify some discussions that are not directly relevant to the present study. Two principal decisions need to be explained here. First, the theory of movement will be based on a Probe-Goal-Agree configuration (see e.g. Chomsky 1995, Pesetsky & Torrego 2007). The main assumption is that lexical items enter the derivation with sets of features, e.g. tense, aspect, or ϕ -features like person and number. These features can at that stage be either valued or unvalued. Unvalued features need to be valued in the derivation, which takes place under an Agree relation: a Probe with unvalued features looks down the derivation for an element, the Goal, with matching features, but crucially a valued version of these features. Under this Agree relation (matching features), the features on the Probe are valued. Agree does not necessarily lead to movement; movement only takes place if there is an additional EPP-feature on the Probe, which stipulates that its specifier must be filled—a further extension of the Extended Projection Principle. The second decision is that Case is taken to be one of the features that need to be valued (following Pesetsky & Torrego 2007; see also Platzack 2005). This means that case is no longer assigned (as in Government & Binding) or checked (as in early Minimalism), and crucially, that the case feature does not by itself drive movement if there is no EPP feature on the Probe. Finally, because one of the questions central to this thesis is that of the possible interaction between syntax and information structure, I will not take this to be a given by including information-structurally motivated projections, as in Rizzi's (1997) approach, in the representations of the clause structures.

Section 1.4 Outline

The aim of this thesis is to unravel the syntactic and information-structural factors involved in the development of the passive and the loss of V2, and to subsequently determine to what extent these factors show an interaction. I argue that the influence of syntax on information structure is evident: syntax limits the options available for information-structural effects, but speakers make the best use of the means that *are* available. In addition, I argue that there is only limited evidence for influence of information structure on syntax; limited, because I find no evidence for

syntactic innovations that are driven by information structure, but despite that, the influence of information structure on the language in general is clear, because it changes the frequency and function of particular constructions. The main arguments of the separate chapters are outlined below.

Chapter 2 provides the starting point for the investigation, by elaborating on the reasons for the hypothesized connection between the loss of V2 and the developments in the passive. In this chapter, I argue that V2 in OE and the passive have two information-structural functions in common: they can rearrange arguments to achieve a given-before-new order of information, and they can place topical information in the preferred position for topics, the initial position in the clause. This chapter also reviews in more detail the proposals by Los (2009, 2012) and Seoane (2006) that the loss of V2—more specifically the importance of the subject as the unmarked theme—led to an increase in the use of passives and other subject-creating strategies. Finally, I argue that in the investigation into the given-before-new function of the passive and V2, two interpretations of givenness need to be taken into account: givenness as discourse-oldness and givenness as cognitive accessibility. The last section of Chapter 2 provides the basics for the annotation of these notions in the historical corpora.

Chapter 3 shows that the introduction of the new passives—the recipient, prepositional and ECM passive—precedes the earliest date for the initial stages for the loss of V2; the new passives are already attested before 1400, albeit marginally. This leads me to conclude that the loss of V2 did not contribute to the introduction of these new passives. Only in the case of the recipient passive can one aspect of loss of V2 be said to be the cause: the recipient passive became possible only after the dative-fronted passive, in turn facilitated by fronting of datives to the topic position in V2 clauses, was lost. I propose that the introduction of the new passives is caused by specific changes in the context for these passives, connecting it to other changes taking place in the transition from OE to ME. Most crucially, however, it is clear that the introduction of these new passives precedes the information-structural consequences of the loss of V2, with the loss of V2 more or less completed only around 1500, over a century after the first attested examples of the three new types of passives.

Chapter 4 shows that in OE V2, three constructions—passives, object fronting and PP preposing—share a function of information-rearranging, while they also have additional individual functions. The passive already functions as an information-rearranger in OE, although to a more limited extent than in later periods. Clause-initial PPs and objects represent alternatives to the given-before-new function of the passive. In addition, clause-initial PPs and objects are shown to have a diverse nature with respect to information-structural behaviour, in the majority of cases containing unmarked ‘local’ links to the discourse, but in a considerable number of examples marked, i.e. contrastive or focused, elements. The analysis of the information status of subjects shows that there is a general tendency for given subjects to precede the finite verb and for new subjects to follow the finite verb, but

that these are only tendencies and cannot be translated into a stricter relation between position and information status. In sum, Chapter 4 shows that the given-before-new principle surfaces in different aspects—subject placement, use of first position—and with different purposes: it either determines variation, as in subject placement, or it is the main motivation for the use of a particular sentence type, as in the choice between passive, object fronting and PP preposing.

Chapter 5 investigates the information-structural aspects of the loss of V2. It adds to the existing body of literature on the loss of V2 by providing details on the loss of inversion in PP-initial clauses, which turns out to be slightly later than in object- and adverb-initial clauses. What also makes the PP-initial clauses stand out is that in EModE, a pattern of late subjects emerges that heralds the so-called locative inversion pattern in PDE. These sentences, often with a locative PP in first position, still show inversion in PDE (*Into the room came a weather man*). The studies in Chapter 5 also show that there is a considerable increase in the proportion of subject-initial clauses in the language, most notably in the transition from EModE to MBE. At the same time that subjects become the default way to begin a sentence, the information-structural character of clause-initial PPs changes, in that they are less often a neutral discourse-linker, and more often contrastive or a discourse-new frame setter. The observations on the aspects of the loss of V2 and the changing character of the clause-initial position lead me to conclude that apart from the loss of one of the subject positions, most of the changes connected to the loss of V2 are in fact related to the information-structural character of the positions, rather than a loss of syntactic possibilities.

Chapter 6 investigates possible consequences of the loss of V2, specifically following the increased importance of the subject as the clause-initial element. It presents a study of the frequency of passives, both of short passives and long passives and of the new passives. These studies all show a general increase in the use of passives, roughly from EModE to MBE. I also discuss middles and non-agentive subjects and show their close structural relation to passives, which lies in their ability to allow an internal argument to surface as the subject. I conclude that this entire range of non-prototypical subject strategies—passives, middles and non-agentive subjects—represents a response to the greater need for subject strategies after the loss of V2. They are all constructions that allow for a non-agentive and often base-generated object to occur in the subject position, thus providing the syntactic means to create those subjects that English needs, but also providing additional options for rearranging arguments, specifically for information-structural aims, in the limited freedom PDE has for reordering.

Chapter 7 provides an answer to the main question of this study: what does the case study of passives and the loss of V2 show about the interaction between syntax and the information-structural principle of given before new? I argue that these case studies do not provide evidence for information-structural factors influencing syntax, in that information-structural factors do not cause syntactic innovations. For instance, the subject strategies as constructions cannot be shown to be introduced

precisely because of V2. Another main indication is that while subject placement in OE follows general given-before-new tendencies, the non-strict relation can only be taken to mean that the placement is not information-structurally driven. What also becomes clear from the studies, however, is that there is a large area of the language where information structure has an important influence on the choice of the syntactic construction. Information-structural needs can, for instance, lead to an increase of specific constructions, such as passives and middles, because of a lack of other means to express a similar information-structural function. They can also completely change the use of one position, such as the change from the versatile first position in information-structural terms to the marked position it becomes in PDE. Ultimately, then, information structure depends on and is limited by what the syntax makes available, but within those limits it has freedom and can change the face of a language entirely.

2. The Information-Rearranging Potential of Passives and Verb Second

Section 2.1 Introduction

Before the corpus studies are presented in the following chapters, this chapter outlines in more detail the reasons behind the hypothesized connection between the loss of verb second (V2) and the developments of the passive in the history of English. I show that V2 and the passive have in common a potential for information-rearranging, with the purpose of either creating an unmarked given-before-new order of information or placing topical information in clause-initial position, the unmarked position for such information. As a result, it is highly probable that syntactic developments in either of these clause types have an influence on the use of the other clause type because of their shared information-structural function.

V2, originally seen as a purely syntactic constraint, has in recent years been readdressed in terms of its information-structural properties, in general and for Old English (OE) in particular. The variability in the OE V2-system—with variation between V2 and verb-third (V3) main clauses—means that a strictly syntactic account cannot do justice to the system as it was in place. Most importantly, it has been proposed that the variation in subject positions is determined by information-structural factors (cf. Van Kemenade & Los 2006; Van Kemenade, Milićev & Baayen 2008). In addition, it has not only been proposed that the system itself is governed by information-structural factors, but also that there are certain properties that

follow from this sensitivity to information structure, such as a specific use of the presubject initial position for discourse-linking (Los 2009).

Like V2, the passive is first and foremost a syntactic phenomenon, but it has a clear function of information-rearranging and reordering arguments in the sentence for discourse cohesion purposes, with proposals describing this function going back at least to the sixties (e.g. Halliday 1968, Chafe 1970). While traditional syntactic accounts have focused on grammatical aspects of the original position of the subject in the passive, functional and discourse linguistics have for a long time focused on a different property of the passive, namely that it can rearrange arguments in the clause, thereby also rearranging the information in the clause. This way the passive can be used to make a proposition fit better into the discourse (cf. Halliday 1968, Siewierska 1984).

What V2 and the passive have in common, then, is that they rearrange arguments in the clause and in that way can be used for information-structural purposes; however, the details of the overlap in function are not evident. Most importantly, while the information-structural function of the passive is reasonably well-defined and has been described in quite some detail, the OE V2 system allows for different types of information-structural functions, for instance not only reordering arguments and placing arguments in clause-initial position, but also variation in other positions in the clause with reference to information-structural principles. This chapter therefore aims to arrive at a clearer view on the type of information-structural functions that have been described in the literature for a range of other constructions, and, since we depend on corpora for historical research, how these information-structural functions can be studied in a corpus.

The chapter is structured as follows. Section 2.2 describes the details of the V2 system as it was used in OE, including the information-structural properties and the syntactic possibilities, followed by an overview of the time and details of the loss of V2 and the proposed information-structural and syntactic consequences of this loss. Section 2.3 turns to the passive, providing a definition of the construction, a description of its functions, and a Minimalist analysis of passive clauses. Section 2.4, finally, zooms in on the literature on information-structural functions, in order to define in more detail the functions that could be applied to describe the overlap between the passive and V2. This section concludes with a discussion of several schemes that have been used to annotate information-structural notions in corpora, both for PDE and OE. Section 2.5 concludes this chapter.

Section 2.2 Old English verb second and its loss: information-structural aspects

This section outlines the information-structural potential of the syntactic properties of the V2 system in OE and the information-structural aspects that are involved in the loss of the system over the course of the Middle English (ME) period. The system in OE is variable, and syntactic factors alone cannot explain all of the attested

variation. In recent years, the focus has therefore shifted from syntactic factors to information-structural factors as an underlying principle to account for the variation. This section first discusses in more detail the technicalities of the V2 constraint in OE and then aims to provide a general time frame for the loss of V2, identifying relevant factors which influence inversion. Finally, this section discusses recent publications—most notably Los (2009, 2012) and Seoane (2006)—which focus on the information-structural properties of the V2 system in OE and the consequences of the loss of the syntactic constraint: an increased use of passives, the introduction of new passives, and an increased use of (relatively) new subject strategies, some of which are likely to have been introduced after the loss of V2.

2.2.1 Categorical and variable verb second in Old English

The V2 constraint, most commonly associated with present-day Germanic languages such as German and Dutch, involves an asymmetry between main and subordinate clauses, and entails that the finite verb always occurs in second position in main clauses, either following a subject in first position, or following an XP—in which case the finite verb itself is followed by the subject. Following Van Kemenade's (1987) proposal for a V2 system in OE, the properties of the constraint in OE have been investigated in detail, and there is by now a substantial body of literature on the irregular nature of the constraint. While earlier attempts at explaining the variation have focused on syntactic factors, such as a possible clitic status of pronouns (Van Kemenade 1987) or variability in the position of the head I within the IP (Pintzuk 1991, 1993), in recent years the attention has shifted to information-structural factors as a governing principle for surface variation in the order of subject and finite verb.

In the literature on OE V2, most authors now make a distinction between two types of V2: categorical and variable. The first, categorical type of V2 is similar to German and Dutch V2 in that it is regular; it occurs in main clauses beginning with an operator, a negator or a discourse adverb like *þa* or *þonne*. In these sentences, illustrated in (1), the finite verb always occurs in second position.

- (1) *þa gemette he ðær ænne þearfan nacodne*
 then met he there a beggar naked
 'Then he met a poor man, naked'
 (ÆLS (Martin) 61—62; Los 2009: 100, her 2)

The analysis of these sentences follows the standard analysis of V2 in present-day languages (e.g. Vikner 1995): the initial element is in Spec,CP and the verb always moves to C, as illustrated in Figures 2.1 and 2.2, based on the modern Dutch examples (2) and (3), which show the alternation between a subject-initial and an XP-initial clause.

- (2) *Ik ben vorige week naar Manchester geweest.*
I am last week to Manchester been.
'I went to Manchester last week'
- (3) *Vorige week ben ik naar Manchester geweest.*
Last week am I to Manchester been.
'Last week I went to Manchester.'

Figure 2.1 Structure of subject-first main clauses

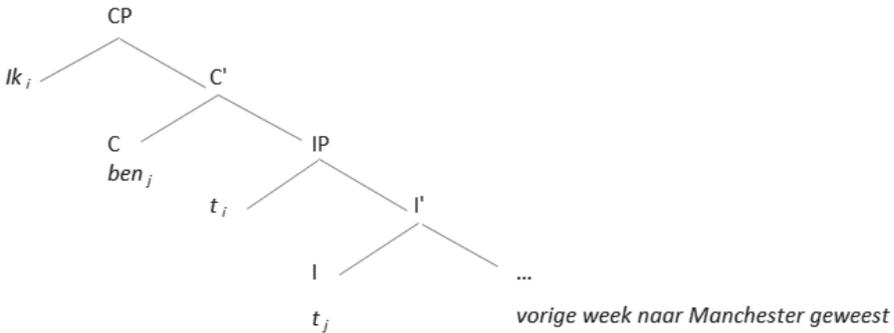
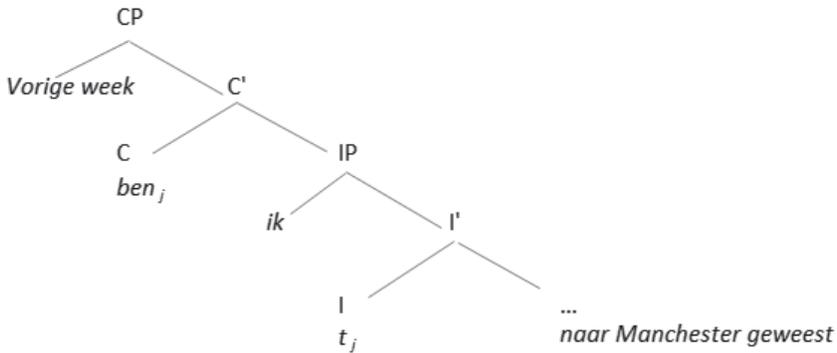


Figure 2.2 Structure of XP-initial main clauses



As in modern V2 languages, the finite verb in clauses beginning with such an operator-like element is assumed to always move to C, with the subject immediately following it, in Spec,IP; in split-IP models (following Pollock 1989), this position is Spec,TP. Note that for ease of representation and to show the similarities between the two sentences, Figure 2.1 presents the subject in the subject-initial clause in Spec,CP and the finite verb in C—as it is in the V2 clauses—but this is not uncontroversial (see e.g. Van Kemenade 1997; Hinterhölzl & Petrova 2010: 321-322 and references cited there): it has both been proposed that subject-initial clauses do

not have a C-domain (Zwart 1997) and that all clauses have a C-domain, in which case the verb always occurs in C, regardless of whether it is preceded by a subject or an XP (Vikner 1995). Part of the discussion is the question whether the Spec,CP position is reserved for topics only, which *may* be subjects, or whether it is always lexicalized, even when—as in subject-initial clauses—movement of subject and finite verb is not visible. Since the focus in this thesis is on possible V2 contexts in clauses not beginning with subjects, this issue need not be resolved here.

The second, variable type of V2 in OE occurs in clauses where the first element is not an operator but a PP, object or adverb. It is in these sentences that variation is found in the order of subject and finite verb, as illustrated in (4) and (5), which is why the focus in recent years has been on this type of sentence.

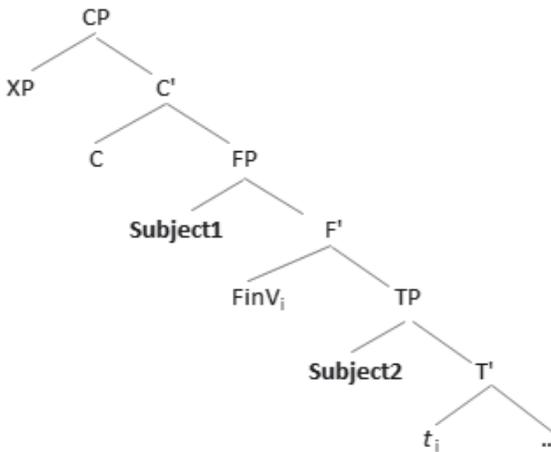
- (4) *Mid þusend searocræftum wolde se swicola deofol þone halgan wer on sume wisan beswican,*
 ‘With a thousand wily arts did the treacherous devil strive in some way to deceive the holy man,’
 (o3, coalive [Martin]:710.6427)

- (5) *On sumere tide Martinus stah to anre upflora,*
 ‘On one occasion Martin was mounting to an upper floor,’
 (o3, coalive [Martin]:601.6358)

Example (4) shows a V2 clause: the initial element—the PP *Mid þusend searocræftum*—is immediately followed by the finite verb *wolde*, which is in turn followed by the subject *se swicola deofol*. Example (5), in contrast, shows a verb-third clause: the clause-initial PP *On sumere tide* is immediately followed by the subject *Martinus*, and only then by the verb *stah*. The by now standard analysis for these sentences is that the verb does not move to the CP domain but only up to the IP domain. In split-IP proposals this position is identified as AgrSP (e.g. Haerberli 2000, 2002, Warner 2007) or FP (e.g. Fischer et al. 2000, Van Kemenade 2012), a functional projection on top of TP.³ This, in turn, leads to a situation with two landing sites for subjects in the IP—either before the finite verb (Spec,FP) or after the finite verb (Spec,TP)—giving rise to the surface V2 or V3 sentences, as illustrated in Figure 2.3.

³ Because object pronouns can also occur in the specifier of this projection (see discussion in Fischer 2000 et al.: 125-126), I will use FP for this projection, and not AgrSP.

Figure 2.3 Subject positions in OE XP-initial main clauses.



During the past three decades several factors have been shown to determine at least to some extent the variation between these subject positions. One such factor is the nature of the subject: pronouns are shown to occur almost invariably before the finite verb while nominal subjects generally follow it (Van Kemenade 1987, Pintzuk 1991). Both Van Kemenade (1987) and Pintzuk (1991, 1993) have proposed that personal pronouns behave like clitics, which are attached to either the left or the right of the verb, although the details of their analyses differ fundamentally. However, it has been demonstrated in later studies that it is not the case that only pronouns occur in verb-third sentences, as already illustrated in example (12) above, although it is true that pronouns hardly ever occur after the finite verb. Another factor which has been shown to determine some of the variation is the type of XP. Even within the class of non-operators there are differences between the extent to which certain types of adverbs ‘trigger’ inversion of subject and finite verb. Warner (2007), for instance, has shown that in ME, *then*, *thus* and *now* almost invariably trigger inversion, comparable to *þa* and *þonne* in OE; crucially these adverbs also trigger inversion with pronominal subjects, indicating that these clauses have the categorical type of V2 rather than the variable type. None of these proposals, however, are able to account for the variation in a principled way.

More recently, proposals have incorporated information-structural aspects to account for the variation found in OE main clauses, based on observations about the distribution of information and specifically the information status of subjects in different positions in the clause. The first study to take the information status of subjects into account in a systematic way is Bech’s (2001) investigation of word order in OE and ME. She presents a study of different word orders and investigates the ‘information value’ of subjects in these sentence types, using the term low

information value, corresponding roughly to ‘discourse-old’, and high information value, corresponding to ‘discourse-new’ (2001: 150-151; 156). Bech finds that preverbal subjects in early OE tend to be old, with 85.2% and 89.8% of the subject being old in non-conjunct clauses and conjunct clauses, respectively (*ibid.* 160). In inversion clauses, subjects tend to be new: 69.5% in non-conjunct clauses and 71.4% in conjunct clauses (*ibid.* 162).⁴

Similar observations about subject placement in V2 main clauses have been made on the basis of studies of the behaviour of subjects in subclauses, by Van Kemenade and Los (2006), Van Kemenade, Milićev & Baayen (2008) and Van Kemenade & Milićev (2012). Following observations that pronominal subjects always occur before the adverbs *þa* and *þonne*, they find evidence that nominal subjects that occur before *þa* and *þonne* in subclauses are likely to be definite (i.e. plausibly given), while subjects occurring after these adverbs tend to be indefinite (i.e. likely to be new). These observations are later taken to generalize to main clauses as well, with Van Kemenade & Westergaard (2012) and Biberauer & Van Kemenade (2011) identifying the finite verb in a similar role as the adverb in separating given from new subjects.

Some researchers have taken these observations as the basis of a formal analysis, most importantly Biberauer & Van Kemenade (2011) and Hinterhölzl & Petrova (2010). Abstracting away from the details of Biberauer & Van Kemenade’s analysis, the crucial aspect is that all subjects, old and new, move to Spec,TP (in their account either on their own or as part of the vP). Only those subjects that represent old information move on to Spec,FP, the higher subject position.⁵ The subjects that remain in Spec,TP position, then, are predicted to be new.⁶ A similar proposal about the information-structural aspects of the high and low subject positions has been made by Hinterhölzl and Petrova (2010), who claim that the finite verb in main clauses acts as an information-structural divider in OE. They propose that in non-operator contexts⁷ the finite verb separates the aboutness topic and “other background elements”—defined as given discourse referents—from the focus material (2010: 319). Their explanation for subject placement in OE, then, is that V2 occurs “by accident”, if there is only one given element in the clause (*ibid.*).

In addition to the two high positions—showing variation between V2 and V3 order—surface inversion can also occur when the subject occurs towards the end of the clause, in a position which is predominantly restricted to verbs which take no

⁴ These are Bech’s (2001) figures for subject-verb clauses when *þa/þonne* and pronominal subjects are excluded, which she unfortunately does not do in all her studies.

⁵ Biberauer & Van Kemenade formalize this by proposing that these DPs have an additional feature that needs to be valued. Interestingly, this is not an information-structural feature, but a morphological one (*gender/person*).

⁶ It is not entirely clear whether the prediction is that they should be new or rather that they cannot be old.

⁷ Hinterhölzl & Petrova do not use this term, but make a distinction between coordinating and subordinating relations, which seems to correlate to the operator and non-operator contrast.

external argument, i.e. unaccusative verbs and passives. The standard analysis for these sentences is that the grammatical subject (the internal argument) remains in its original position inside the VP (cf. Van Kemenade 1997, Warner 2007). Example (6) illustrates an inverted subject and (7) a so-called ‘late subject’ (Warner 2007).

(6) *And þen seid þe deuelles to ham . . .*

and then said the devils to them

‘And then the devils said to them . . .’

(RevPurg, line 534; Warner 2007: 85, his 3b)

(7) *þanne wente out to lon þe puple of Ierusalem*

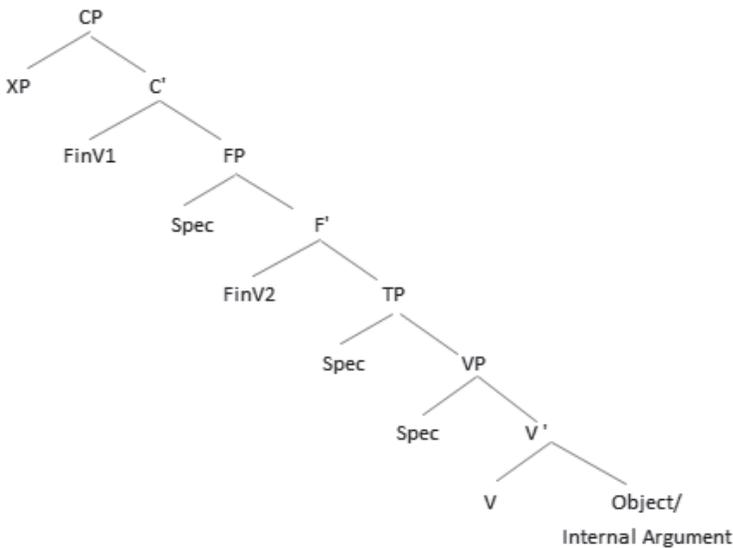
then went out to John the people of Jerusalem

‘Then the people of Jerusalem went out to John’

(Wyclifite Sermons, vol. III 124.29; Warner 2007: 85, his 3cii)

While both inverted subjects and late subjects occur after the finite verb, only inverted subjects occur in the Spec,FP position immediately after the finite verb; in these examples the additional material in the VP—the prepositional phrases *to ham* and *to lon*—works as a diagnostic to determine the position of the subject as inside or outside of the VP. Figure 2.4 illustrates the structural analysis of these late subjects, with the subject in its original internal argument position inside the VP.

Figure 2.4 Position of late subjects in OE main clauses



Like the inverted subjects, these late subjects are likely to be new, as shown by Warner (2007: 93-96) for ME: he provides evidence that indefinite subjects and

heavy subjects (measured by length) occur more often in the late subject position than in either of the two higher subject positions and points out that indefiniteness and weight have a strong correlation with new information. Biberauer & Van Kemenade (2011: 34), for OE, also predict that these late subjects are new: old and negative subjects are raised out of the VP, and so stranded subjects must be old.

In short, V2 in OE comes in two types—categorical/V-to-C and variable/V-to-F—and the variable type of V2 allows for three subject positions—high, inverted, and late—which have been associated with particular information-structural properties: high subjects are expected to be old, while inverted and late subjects are expected to be new. With respect to the structure of the main clauses and the variation between V2 and verb-third main clauses, two questions remain unresolved for now. The first question is to what extent the information status of the subject determines the variation between the two subject positions. The second question is a theoretical one: based on the strength of the correlation, can subject placement be said to be formally driven by information-structural principles?

2.2.2 A time frame for the loss of verb second

It is generally assumed that V2 is lost during the fifteenth century (Van Kemenade 1987, Fischer et al. 2000: 129-137), but, as various studies have shown, the timing of this loss is not straightforward for a number of reasons: the system was variable in OE to begin with; inversion does not decline at the same rate in all contexts; and finally, not all types of inversion are lost over the course of the centuries. This section presents an overview of studies tracing the loss of V2 in different contexts, focusing on the V-to-F contexts, i.e. the type of V2 that was variable in OE. I focus on the studies that follow the analysis of V2 presented above instead of only incorporating surface patterns. As Fischer et al. (2000), Los (2009) and Warner (2007) point out, failure by scholars to take these fundamental structural aspects into account is one of the reasons for the lack of a coherent account in the literature for the exact time of the loss of V2.

The first type of V2 that was identified in the previous section, that which occurs in negative contexts and in questions—with the initial element in Spec,CP, the verb in C and the subject in Spec,TP—is in fact not lost: it survives into PDE as question inversion (8) and subject-auxiliary inversion (9).

(8) When did this happen?

(9) Rarely did I hear such overtones of gratitude as went into the utterance of this compound noun. [=Green 1980, ex. 32e]

(Birner & Ward 1998: 157, their 199)

What has changed, however, is that V-to-I movement is lost and as a consequence this type of inversion in PDE only occurs with auxiliaries and dummy *do*, and no longer with lexical verbs, as was the case in OE. As Fischer et al. (2000: 134) argue,

the loss of V-to-I is a separate development from the loss of V2, i.e. the loss of V-to-C or V-to-F movement; it should therefore not concern us here. A special case within the group of V-to-C type of inversion in OE is the *þa/þonne* group, which, in contrast to negative-initial clauses and questions, no longer trigger inversion in PDE (cf. **Then was he knighted*). Haeberli (2007: 23, for ME) and Van Kemenade (2012: 830-831, for ME and EModE) provide data on the loss of inversion after *then* and other members of the *then*-group: the rates of inversion already decrease throughout the ME period and inversion is almost completely lost by the end of the EModE period. The loss of this specific type of inversion must be seen as a separate, specific development within the loss of V2 and will therefore not be considered further.

The second, variable, type of V2—with a non-operator in Spec,CP, a nominal subject in Spec,FP and the finite verb in F—is lost at different speeds in different contexts. Fischer et al. (2000: 132-133), summarizing findings from Jacobsson (1951) and Schmidt (1980), provide the following timing: a “sharp decline” in the late fourteenth and early fifteenth century, variability in the sixteenth century, and a final loss in the seventeenth century. Since then, several studies have been published which present increasingly detailed quantitative data—something which has become possible because of the availability of the syntactically annotated corpora (the PPCME, for instance, became available in 2000; the YCOE in 2003). It is important to note that inversion, in the earliest periods, is not the only possible word order, as the discussion on the variability of this pattern has already made clear. For OE, Haeberli (2002) reports rates of inversion between 70% and 75% (in text dated before 950 and after 950, respectively). Within this general rate of inversion, there are considerable differences between individual texts, ranging from 40.5% (Gregory C) to 84.8% (Apollonius) (2002: 250). These widely varying rates persist for a number of centuries: Kroch & Taylor’s (1997) provide similar percentages for early ME, with an overall rate of inversion of 71.4%, which is within the range provided by Haeberli for OE (average percentage provided by Haeberli 2002).

It is towards the end of the fourteenth century that the situation starts to change. Haeberli (2002) shows that the rates of inversion in the late fourteenth and fifteenth century have become considerably lower, and he characterizes inversion as a “clear minority pattern” in this period: inversion is down to less than 50% in most of his text samples (2002: 254-5), while it is lower than 25% in roughly half of his selected texts.⁸ Another general overview of the rates of inversion in the fourteenth and fifteenth century is presented by Warner (2007)⁹, who reports an overall rate of inversion of 40%. Like Haeberli, Warner notes that there is considerable variation between individual texts (2007: 48).

Variation not only exists between different texts—as Warner and Haeberli both find—but also between different types of contexts. First, several adverbs have been

⁸ Haeberli only selects non-pronominal subjects with non-operators in initial position, and excludes unaccusatives and clause-initial *þa*, *þonne* and *nu*.

⁹ His study is based on 120 main declarative clauses from 32 texts (Warner 2007: 47).

identified which behave similarly to *pa/ponne* in that they trigger inversion almost invariably; Warner calls these the *then-group*, consisting of *then*, *thus* and *now*. Second, the syntactic category of the initial element influences inversion. Van Kemenade & Westergaard (2012) investigate inversion in ME in sentences with adverbs and objects in initial position, which they identify as information-structural links to the preceding discourse—adverbs as “discourse continuation” and objects as marked, often contrastive, links (2012: 103). Their results show that inversion in these two types of clauses declines at a different rate: while inversion—in a selection of clauses with transitive/unergative verbs and nominal subjects—after adverbs is down to just over 20%, inversion after objects is still at 82.5% in M4 (although this is based on a small number of examples). Finally, variation has been found with respect to the syntactic status of the clause-initial element: Warner (2007: 99) finds that it makes a difference in his data for ME, the period of change, whether the clause-initial element is an argument of the verb or an adjunct: arguments occur more often with inversion.

In addition to the type of element in first position, the type of verb has also been shown to influence inversion. Early studies, such as Van Kemenade (1997) and Haerberli (2002), single out the category of unaccusatives, which show higher rates of inversion than other types of verbs. Van Kemenade & Westergaard (2012) attempt to determine in more detail the effect of the type of verb, distinguishing between unaccusative main verbs, auxiliaries and transitive/unergative verbs. They provide substantial evidence confirming earlier observations that inversion is lost at a lower speed with unaccusative verbs: while inversion with nominal subjects and transitive/unergative verbs declines in ME (from 66.8% in M1 to 22.9% in M4¹⁰), inversion with auxiliaries and unaccusatives remains high (48.9% and 57.2%, respectively). Van Kemenade (2012) provides a follow-up to this study, narrowing down the focus to adverb-initial clauses but extending the period to include EModE. She finds that in these adverb-initial clauses, transitive/unergative and auxiliary verbs show a steady decline of inversion from M1 onwards, while unaccusative verbs continue to occur with inversion. Most importantly, Van Kemenade’s data show that subjects of unaccusatives continue to occur after the finite verb even when inversion with transitive/unergative and auxiliary verbs is lost already. In E3—roughly the second half of seventeenth century—inversion with transitive/unergative and auxiliaries is down to 15.6% and 12.7%, respectively; at the same time, however, unaccusatives still show inversion in 38.3% of clauses beginning with an adverb which is not in the *then-group* (2012: 829).

One reason why unaccusatives are particularly relevant is that these are the verbs that can occur with late subjects, i.e. the internal argument in these clauses can remain inside the VP. Unlike the rest of the V-to-F category, this type of inversion

¹⁰ For these particular numbers, Van Kemenade & Westergaard (2012) do not distinguish between types of initial constituent.

is not lost completely: it survives partly in PDE as locative inversion, illustrated in (10).

- (10) There are huge cartons and tins of nuts, vanilla, honey, peanut butter. Varieties of herbs tea are visible. *On the counter are loaves—whole wheat, cinnamon raisin, oatmeal, rye, soy sunflower, corn meal.* [Terkel 1974:607]
(Birner & Ward 1998: 156, their 198a)

This type of inversion pattern in PDE is different from the subject-auxiliary inversion and question inversion in (10), in that it is an optional pattern. As in OE and ME, it occurs predominantly with unaccusative verbs, and in PDE it almost exclusively occurs with locative PPs—something which does not seem to have been the case in OE, and something for which the precise motivation and development is as yet unclear. With respect to the decline of this pattern, it is important to note that not all unaccusative subjects are necessarily late subjects in earlier English (cf. Van Kemenade & Westergaard 2012: 102). Consider example (11), which illustrates the ambiguity in structure for late and immediately postverbal subjects.

- (11) *Of þese seuene heuedes comen alle manere of synnes*
from these seven heads come all manner of sins
'From these seven heads spring all manner of sins'
(Vices and Vertues 11.8; Warner 2007: 94, his 11)

As Warner (2007: 93) points out, it is not possible to tell whether the subject is in immediately postverbal position (the V2 position, i.e. Spec,TP) or whether it has remained low, within the VP. These late subjects seem to have acquired a separate function of introducing new (or indefinite and long) information, and their development is likely to be different from that of the inverted subjects, as indeed the data on unaccusatives already suggest. It is not the case, however, that the information-structural function for this type of subject is reserved for subjects of unaccusatives, since, as Warner (2007) shows, subjects which occur in clause-final position in clauses with other types of verbs are also often new and/or long (2007: 97f). The following examples illustrate the difference between the clause-final position (12) and what Warner calls the "internal context", as in (13).

- (12) *Alswa, clenness of mouth, kepes þre thynges: Ane es þat þou umthynk þe before, or þou speke.*
also, cleanness of mouth preserve three things one is that you consider you before ere you speak
'Also three things preserve purity of mouth. One is that you reflect before you speak.'
(Rolle, The Form of Living, 25; Warner 2007: 98, his 13b)

- (13) & þus gat *kyng acab* the vinezard.
 and thus got King Ahab the vineyard
 ‘And thus King Ahab obtained the vineyard.’
 (Lavenham 9.37; Warner 2007: 99, his 14b)

The subject in (12) occurs in clause-final context; what follows is a new clause. The subject in (13), in contrast, occurs clause-internally: it is followed by an object. The most straightforward analysis of subjects as in (13) is that they are not situated in the low position inside the VP, but in Spec,TP, meaning that they do not only occur with unaccusative verbs, but can occur with all types of verbs. In other words, while unaccusatives provide the option of clause-final subjects most straightforwardly because they allow for a subject to remain within the VP, this information-structural function of presenting new information in final position is—within the word order options in OE and ME—available for subjects of other verbs as well, in those cases where there is no additional material inside the VP.

In conclusion, what these corpus studies on specific V2 contexts in OE show is that the major change from inversion as a frequent pattern to a minority pattern takes place between 1400 and 1500. What also becomes clear, however, is that there are considerable differences between contexts, both at the time that V2 was still in place as well as in the decline of the inversion pattern. Two factors seem to be especially relevant: first, the type of verb, with unaccusatives occurring with inversion more frequently and for a longer time than other types of verbs; and second, the clause-initial element: not only are there the operator-like elements which almost invariably trigger inversion, but among the class of non-operators there are certain elements which trigger inversion more often than others, such as arguments, certain adverbs and locative PPs.

2.2.3 The pragmatic consequences of the loss of verb second

The new view on the nature of V2 in OE and the information-structural principles underlying subject placement in OE has in recent years also changed the ideas on what the loss of such a system means for the language. This specifically concerns the environments in which there is variation between V2 and V3 sentences, i.e. non-operator initial clauses with V-to-F movement. In particular, it seems that the V2 system allows for more word order options and brings with it a specific use of the clause-initial position, a view that is based on observations on both old and present-day V2 languages (see e.g. Los 2009, Hinterhölzl & Petrova 2010). The main consequence of such a view is that the loss of the syntactic constraint on verb placement is likely to have had both syntactic and pragmatic consequences for the language.

Los (2005, 2009, 2012) proposes that the loss of V2 is not only a matter of syntactic changes, but that the syntactic changes are accompanied by pragmatic changes in the use of the subject and the pre-subject position. Central to Los's proposal is the idea that there is a change from OE to PDE in the encoding of

unmarked themes, a Functional Grammar term, which Los—referring to Downing & Locke (1995) and Halliday (1994 [1985])—explains as being used to denote “old information with [a] lack of prominence” (2009: 113). Languages can differ in the means that are available for encoding unmarked themes, which Los illustrates with the following example of a translation of a Dutch sentence from Hannay & Keizer (1993).

- (14) *En daarmee was de tragedie van Bergkamp compleet*
and with.that was the tragedy of Bergkamp complete
'and that made Bergkamp's tragedy complete'

(Los 2009: 113, her 20)

Los explains how a literal translation of *en daarmee* with *and with that* would be syntactically possible in PDE but how this would place more prominence on this phrase than it has in the original Dutch sentence: in PDE *with that* is a marked theme while *daarmee* in Dutch in an unmarked theme. The only way to translate *daarmee* as the unmarked theme is to make it the subject of the English sentence, as in *that made Bergkamp's tragedy complete*. Los proposes that the distribution of unmarked themes has changed from OE to PDE: where OE, much like present-day Dutch, had two positions available for unmarked themes—the subject and the presubject clause-initial position—PDE only has one, the subject position.

What stands out most about the presubject position in OE in comparison to PDE is its versatility: in addition to unmarked themes, it can—using a variety of syntactic categories—also express marked themes, the type that remains available in PDE. The following two examples illustrate how both the subject and the element in presubject position can be unmarked themes.

- (15) *Be þære he gestrynde dry suna Her & Onam & Sela.*

By that he begot three sons Er & Onan & Shelah

'He had three sons with her: Er, Onan and Shelah.'

(cogenesiC,Gen_[Ker]: 38.3.83; Dreschler & Hebing 2011: 59, their 5)

- (16) *After þysum dadum hi tancodon Drihtne.*

After these deeds they thanked Lord

'After these deeds, they thanked the Lord.'

(coaelive [Maccabees]:504.5170; Dreschler & Hebing 2011: 59, their 7)

In (15) and (16), both the clause-initial PP and the subject represent discourse-old information and thus both provide a link to the preceding discourse in an information-structurally unmarked way. Example (17) shows that a fronted object—associated with focus in PDE—could be an unmarked theme in OE, providing a similar link to the previous sentence.

- (17) *þone* asende se Sunu,
 this sent the Son
 ‘The son sent this one’
 (coaelhom,ÆHom_9:114.1350; Speyer 2010: 38, his 17a)

þone is discourse-old information and there are no indications in the text that it has a contrastive or focused nature (Speyer 2010: 38). What this means is that object fronting and PP preposing can be used to place old information in initial position, and crucially, to place old information before any new information in the sentence. At the same time, the presubject position in OE could do more than only express unmarked themes: marked themes, such as contrastive or focused elements are also already used, as illustrated in (18).

- (18) ‘Then went he through all the Egyptian land, sowing God’s seed, and healed the sick.’
Hreoflige he *geclænsode fram ðære unclænan coðe*, *Wode* he *gehælde*
 lepers he cleansed from the unclean disease, possessed he healed
 ‘He cleansed the lepers from the unclean disease; he healed the possessed’
 (o3, coaelive [Mark]:6.3189)

The clause-initial objects *Hreoflige* and *Wode* are contrasted with *the sick* in the preceding sentence and with each other.

In the periods after OE, the presubject position remains available syntactically, but becomes information-structurally marked: in PDE it can no longer express unmarked themes. As Los & Dreschler (2012: 864) point out, corpus studies and modern grammars point out that presubject place adverbials are marked in PDE and have become relatively infrequent (e.g. Biber et al. 1999: 803; Breivik & Swan 1994). In addition to being infrequently used, these adverbials are often used contrastively or as frame setters, as in (19).

- (19) A: How is business going for Daimler-Chrysler?¹¹
 B: [*In GERMANY*]_{Frame} the prospects are [GOOD]_F,
 but [*in AMERICA*]_{Frame} they are [losing MONEY]_F.
 (Los 2012: 26, her 8; from Krifka 2007: 45)

Los & Dreschler (2012: 864) provide suggestive diachronic data for this development. Their search in the syntactically annotated corpora of historical English (see Section 1.3) shows that subjects become more frequent in clause-initial position at the expense of other elements: the percentage of subjects in initial position increases from 44% in OE to 65% in Late Modern English while the percentages of objects and adverbials decrease from 4% to almost 0% and 35% to

¹¹ In this example, the small caps indicate accent and F stands for focus.

20%, respectively. Interestingly, this change in markedness is a consequence which does not directly follow from the *syntactic* changes, because a presubject position is still available in PDE.

Another piece of evidence for the observation that the changes in the use of the initial position are not purely syntactic is that a functional change takes place which is not related to the syntactic status of the initial element—the loss of deictic referencing. Evidence from Bohnacker & Rosén (2007) on Swedish and German, as reported by Los (2009), shows that these two V2 languages use the initial, presubject, position differently: German makes more use of deictic referencing in first position. It seems that OE behaves like German in this respect and also uses the first position for deictic reference, as Los (2009) shows, on the basis of a comparison of an OE and a PDE translation of Genesis (2009: 106). The use of the first position changes in later centuries, however, and Los (2012: 37-41) argues how the impoverishment in the demonstrative system, such as the loss of gender, is central to the loss of deictic referencing, in that it meant a loss of possibilities to use a demonstrative for a specific reference. Evidence for the decrease in deictic referencing is provided by Los & Dreschler (2012: 866) and Komen et al. (2014: 103): PPs containing pronominal and demonstrative elements become considerably less frequent in initial position from OE to LModE. This means that the first position lost a functionality which is connected to the morphological system, but which cannot be explained with reference to syntax only: deictic referencing is still possible to some extent in PDE, but seems to be used only infrequently; crucially, singular independent demonstratives can no longer refer to human referents, as they can in OE.

The first position loses another function after the OE period, that of local anchoring: providing a referential link to the preceding discourse. Hinterhölzl & Petrova (2010) propose that the verb in OE separates given from new material—as can be seen in examples (15) and (16) above. Los & Dreschler (2012) similarly point out that a frequent function of this preverbal material is to provide a link to the immediately preceding discourse, which they call ‘local anchoring’. This local anchor can be expressed syntactically in different ways: as subject, object or adverbial. Komen et al. (2014) show that the function of local anchoring is lost in the following century, in their study of the referential character of clause-initial PPs in a selection of main clauses. Their results show a clear decrease in the proportion of referential PPs from ME to Late Modern English, from 58% to 22%.¹²

There are also indications that the function and use of the subject itself change under the influence of the developments in the presubject position. Los & Dreschler (2012: 868-9) point out some of these aspects on the basis of a single text, comparing an OE and PDE version of *Genesis*: the percentage of subject-initial

¹² Komen et al. (2014) find a (statistically not significant) increase from OE to ME; the corpus is small (in this case, 3 OE texts and 4 ME texts), which means that the number of examples overall is low (ranging from 22 to 124 for the periods).

clauses is higher in PDE than in OE (75% against 55%); and the percentage of ellipted is lower in PDE than in OE (7% against 17%). Komen et al. (2014) provide further evidence for this. They aim to find evidence for the hypothesis that the subject takes on more functions after the versatile presubject position is lost. What they predict is that the subject becomes less ‘stable’ because it takes over some of the functions from the OE presubject position: while the presubject position cannot express unmarked themes any longer, the need to express unmarked themes remains, in addition to the functions the subject already expressed. This increased the functional load of the subject. This would mean that the subject will be increasingly likely to refer to a different referent from one sentence to the next—instead of for instance only encoding the protagonist in a story. The results of their studies suggest that this is exactly what happens after ME. First, subjects are less often ellipted in PDE than in OE, which suggests they are less predictable. Second, subjects are more often inanimate, which can be explained by their new function as encoding local anchors, which are unlikely to be animate. Finally, they show that subjects more often change: in the selected PDE texts the subject of a sentence is more often different from the subject in the previous sentence than in the selected OE texts (roughly 55% in OE against just over 70% in ME, EModE and late ModE).

The consequences of the loss of V2 in terms of the use of the first position and the changing function of the subject in turn led to other changes in the language. Los (2009) concludes that the increasingly marked status of the presubject position leads to “a greater need” for subjects, because the subject is now the only way to start a sentence in an information-structurally unmarked manner (2009: 118). This follows from an assumption that the need for unmarked sentence beginnings is likely to have stayed as strong over the periods. The way the language responds to this, she argues, is through an increase in subject strategies, i.e. ways to put an argument in subject position. Most specifically, Los points to the introduction and establishment of new passives: the ECM passive, the prepositional passive and the recipient passive—which were all introduced around the time that verb second was lost; ECM passives especially, according to Los, are motivated by information-structural needs. Another strategy pointed out by Los (2009) is the use in English of a range of non-agentive subjects, which Rohdenburg (1974) has shown to be ungrammatical in German.

- (20) a. 2004 saw the advent of routers with turbo modes.
 b. In 2004 kwamen de eerste routers met turbo modes op de markt.
 c. *2004 zag de komst van de eerste routers met turbo modes.
 (after www.cnet.com.au/wireless/routers/; Los 2009: 118: her 33)

A final strategy that can also be placed in this category are middles or middle-like sentences—as in *Bureaucrats bribe easily* or *This TV needs fixing*—which have been described as another innovation in the ME and EModE periods by Hundt (2007) and Toyota (2008).

Another perspective on the scenario of the increase in subject strategies is provided by Seoane (2006), who connects earlier studies about the frequency of the passive and the use of the passive as an information-rearranger to the proposal by Los. Seoane (2000, 2006) shows how the long passive is already used in Early Modern English as it is in PDE: it places topical information in subject position, and places given information before new information. In addition, she finds an increase in the frequency of passives with respect to active transitive clauses, from 18.0% in M4, to 24.1% in E3 (2006: 371). In Seoane (2006), she explains this increase as another consequence of the loss of V2, i.e. the need for strategies that can place topical information in subject position.

In conclusion, the relevant information-structural change that takes place in conjunction with the loss of V2 is the loss of the presubject position as a possible encoding of unmarked themes, with the result that the subject becomes the only unmarked theme. As a consequence of this, there is an increased need for subject strategies, as well as a greater functional load for the subject. This development results in new, or an increased use of existing, subject strategies such as passives in general and ECM passives, prepositional passives and recipient passives in particular. In addition, there is a range of non-agentive and inanimate subjects that English allows which also seem to function exactly to fill the gap created by the loss of the presubject position for unmarked themes.

2.2.4 Summary and conclusions

This section has provided an overview of the V2 system as it was in place in OE, focusing on the two different types of V2 that are found in OE: a categorical and a variable type, where the variable type has been said to be largely determined by information-structural factors. It is this second type of V2 that is relevant to the introduction of the passives and an increase in use of passives and subject strategies as proposed by Los (2009, 2012) and Seoane (2006). In addition, the V2 system as it was in place in OE allowed for more word order variation, such as object fronting and PP preposing, which could be used by speakers for at least two information-structural effects: (i) rearranging of arguments to create given-before-new order of information; and (ii) placing topical information in clause-initial position, the unmarked position for topical information. This section has also provided a general time frame for the loss of V-to-F movement and the corresponding V2 word orders, beginning from around 1400, with different contexts showing a different rate of the loss of inversion—most notably, type of verb and type of initial element. Finally, I have reviewed in more detail the proposals by Los and Seoane that the additional positions that the V2 system made available to express unmarked themes was lost over the course of the centuries and that the subject became the only way to express unmarked themes. As a consequence, the language is likely to have looked for new opportunities to create these unmarked theme-subjects and to have found it in an increased use of passives, new types of passives and non-agentive subjects.

Section 2.3 Syntactic and functional properties of the English passive

Like the word order options made available by V2, the syntax of the passive provides a potential for information-structural functions, most importantly rearranging information in the sentence to achieve a given-before-new order of information, as well as complying with general tendencies for subjects to express topical information. At the same time, however, the passive may also serve other functions, such as demoting the agent. This section provides a definition of the passive, and explains the information-rearranging function and the subject-creating function of the passive in more detail. In addition, I present a Minimalist analysis of the passive.

2.3.1 A definition of the passive

Passives are a quite common feature of the languages of the world, but their properties vary from language to language, and discussions in the literature have shown that providing a definition of the passive is not straightforward (e.g. Siewierska 1984, Wanner 2009). In addition, a discussion has taken place about the function of the passive, with different defining aspects being considered as the main function by various authors. This section presents both the formal definition and the main function of the passive (short and long) that forms the basis for the studies in the following chapters.

Characteristics of the passive

As a starting point for a definition of the passive, let us first review the basic descriptive properties of the passive in English. The passive is usually defined in comparison to its active counterpart—in part due to theoretical considerations, especially in early Generative theory, that the active is the basic form (cf. Chomsky 1957). Example (21)a illustrates an active sentence and example (21)b a passive with the same propositional content.

- (21) a. Pat stole my surfboard.
 b. My surfboard was stolen by Pat.

(Huddleston & Pullum 2002: 1427, their 3i-ii)

The link between active and passive is intuitively clear: even though an argument is missing in (21)b, we know that both sentences refer to the same verb *steal*, with the same basic argument structure. On the basis of this example, Huddleston & Pullum provide the following list of syntactic differences between the active and the passive:

- (22) a. The subject of the active (*Pat*) appears in the passive as the complement of the preposition *by*.
 b. The object of the active (*my surfboard*) appears as the subject of the passive.

c. The verb of the active appears in the passive in the past participle form (*stolen*).

d. The passive contains an extra verb, the auxiliary *be*.

(*ibid.* 2002: 1428, their 4i-iv)

The *by*-phrase can of course be omitted; and PDE has developed a second passive auxiliary in *get*, which is usually characterized as a less formal option, and is often used with dynamic, as opposed to stative, verbs (cf. Huddleston & Pullum 2002: 1440-1443).

While these descriptions may seem limited in their ability to capture the essential properties of the passive, cross-linguistic work shows that it is in fact difficult to provide a universal definition of the passive beyond its properties in specific languages. First of all, the passive does not occur in all languages: less than half (43.4%) of the languages included in the *World Atlas of Language Structures Online* make use of passive constructions (Siewierska 2013). Among the languages that do have a passive, there are considerable differences in the properties of these constructions, as Siewierska (1984, 2013) has shown. She gives the following properties as being shared by passives in all languages:

- (23) a. They have a subject which corresponds to a non-subject in the active;
b. The action or event expressed by the verb is brought about by some person or thing which is not the subject;
c. The person or thing if not overt is at least strongly implied and can function as a potential subject of the active.

(Siewierska 1984: 79)

These defining characteristics build on a relation between the passive and active form of the sentence, stating essentially—as in Huddleston & Pullum’s (2002) description—that the active object becomes the passive subject and the active subject becomes an adjunct in the passive sentence. What this list of characteristics adds, however, is the notion of the ‘implied actor’. Indeed, this is how Cornelis (1997) ultimately defines the passive in Dutch—which because of the close relation between the two languages can also shed light on the passive in English: “process towards a final state, the agent of which should not be identified with” (95). Note that these types of definitions already move towards the function of the passive rather than only describing the syntactic properties.

Both the description by Huddleston & Pullum (2002) and the definition from Siewierska (1984) build on quite a close relation between active and passive, and although this relation is not entirely straightforward, it does provide a clear reference point in terms of thematic properties of verbs. While Siewierska (1984: 30-31) points out that the propositional content of an active and passive with the same arguments is generally the same, the passive—especially the short passive—

carries an ambiguity in meaning that is not present in the active, as illustrated in the following examples.

- (24) a. Janet was killed.
 b. Someone killed Janet.
 c. Some people killed Janet.
 d. Lightning killed Janet.
 e. Boredom killed Janet.

(Siewierska 1984:30, her 7a-e)

As Siewierska explains, the short passive in (24)a does not specify who killed Janet, and (24)b-d present several options which in the absence of any further information are all possible. Another example against a direct correlation between active and passive is the observation that not all active transitive clauses can be passivized, even when a fundamental requirement of transitivity is fulfilled, as illustrated in (25)a and (25)b.

- (25) a. A strange adventure befell him.
 b. *He was befallen by a strange adventure.

(Huddleston & Pullum 2002: 1432, their 16ia-b)

In addition, not all passive clauses have an active counterpart, meaning that it can be problematic to say that all passives are derived from an active form. Short passives do not have an “exact” active counterpart (Huddleston & Pullum 2002: 1428), and many ECM passives do not have an active version, as the following examples show.¹³

- (26) Pat is reputed to be very rich.

- (27) a. Kim is said to be a manic depressive.
 b. *They say Kim to be a manic depressive

(*ibid.* 1435-1436, their 31i,iii and no number)

The verb *repute* does not have an active form, and, as (27)b shows, there is also no active equivalent of the use of *say* as in (27)a. Despite these exceptions, the relation between active and passive provides, in general, a useful point of departure for understanding the properties of passive clauses, especially selectional requirements and thematic roles.

¹³ Huddleston & Pullum do not use the term ‘ECM passive’ but simply say that there is a group of verbs with a special complementation pattern.

The function of the passive

The most remarkable of the syntactic properties of the passive is the fact that the agent is either demoted from subject to adjunct status (in long passives) or completely excluded (in short passives). Considering that, according to several thematic or semantic hierarchies, an agent is the most likely candidate to become the subject of a clause (cf. Givón 1984, Siewierska 1984: 220), there must be a strong motivation for omitting it. The exclusion of the agent necessarily has the effect that the patient is promoted, because in a regular transitive clause, it is the only remaining argument of the verb. Both aspects—agent demotion and patient promotion—have been described as the most important reason for using the passive. In this section I argue that the primary motivation for using the passive is the promotion of the patient because of its relevance for placing topical information in the unmarked position for topics, thereby facilitating cohesion in the discourse.

Let us first consider some examples that illustrate the function and cohesive properties of the passive in a text. The examples were taken from a Wikipedia entry on J.D. Salinger. The overarching topic of such an encyclopaedic text is stable throughout, since everything in the text should contribute to the reader's knowledge about this particular person. This means that in principle this person is the topic of each sentence and is expected to be the subject of active sentences when aspects of his life are discussed. This is illustrated in (28)a-c, which gives the first three sentences of the entry.

- (28) a. Jerome David Salinger [...] was an American writer who won acclaim early in life.
b. He led a very private life for more than a half-century.
c. He published his final original work in 1965 and gave his last interview in 1980.
(en.wikipedia.org/wiki/J._D._Salinger, accessed 15 August 2014)

Salinger is the subject of each of these sentences, and these sentences add information about him: who he was and what he did. All verbs are active: *was*, *won*, *led*, *published* and *gave*. Nevertheless, people are not always the agents or actors in their own lives and they may be affected by actions of other people or by outside forces. In these cases, an active sentence with the agent as the subject may still be possible, but a far better choice is to use a passive with the patient as the subject. The following examples illustrate three passive choices in the text on Salinger (in bold), given with the context in which they occur; the subjects and their antecedents are underlined.

- (29) a. As a youth, Salinger attended public schools on the West Side of Manhattan.
b. Then in 1932, the family moved to Park Avenue, and **Salinger was enrolled** at the McBurney School, a nearby private school.

- (30) a. The same year, Salinger began submitting short stories to *The New Yorker*.
 b. **Seven of Salinger's stories were rejected** by the magazine that year, including "Lunch for Three", "Monologue for a Watery Highball", and "I Went to School with Adolf Hitler".
 c. In December 1941, however, the publication accepted "Slight Rebellion off Madison", a Manhattan-set story about a disaffected teenager named Holden Caulfield with "pre-war jitters".
 d. When Japan carried out the attack on Pearl Harbor that month, **the story was rendered "unpublishable"**;
 e. it did not appear in the magazine until 1946.
 (en.wikipedia.org/wiki/J._D._Salinger, accessed 15 August 2014)

In (29)a Salinger is the subject of an active clause, and in the second clause in (29)b, Salinger is retained as the subject of the passive clause; Salinger is clearly the topic of the sentence, while it is irrelevant who enrolled him. In (30)a, Salinger is again the subject of an active clause, and the object *short stories* is new. This object then becomes the subject and topic of the following clause—*seven of Salinger's stories*—where it is the patient of the passive verb phrase *were rejected*. A similar shift takes place in (30)c and (30)d: the object "*Slight Rebellion*"... "*pre-war jitters*" becomes the subject of the passive main clause in the following sentence. In short, the choice for the passive in (29)b, (30)b and (30)d allows for either retaining the overall topic of the text or the continuation of a topic that is introduced in the preceding clause or sentence.

This function of 'foregrounding' the patient has been described in various ways in the literature, in many cases formulated in terms of the information-structural notions of topic, theme, or givenness (cf. Halliday 1967, Thompson 1987). An example of such a definition is presented by Siewierska (1984), who describes the function of the passive as 'topicalizing', defined as the "assignment of subject/topic or just topic function to a non-agent" (1984: 217). 'Topic' is used by Siewierska with two different meanings which both determine the use of the passive: (i) topic as the initial element in the clause; and (ii) topic as given information. Taking these two definitions, the passive either places a non-agent in initial position—a comparable function to fronting strategies¹⁴—or it places a given referent in the position for given information, i.e. in initial position. Crucially, Siewierska defines this second function as a relative constraint: the given information does not just occur in initial position, but it precedes the new information that is expressed by the agent *by-*

¹⁴ Keenan & Dryer (2006), for instance, compare the function of passives to that of topicalization (*Beans I like*) and left-dislocation (*As for the president, congressmen don't respect him any more*), describing it as a "foregrounding" device. They conclude, however, that the foregrounding function of the passive seems to be weaker than it is for topicalization and left-dislocation, and that there is not really a competition between the passive and these constructions.

phrase. Another aspect of the function of the passive described by Siewierska is the effect of using a passive, especially a short passive, in the discourse, namely, “facilitat[ing] the linking of successive clauses”, especially in written texts (1984: 230). As the discussion of examples (29) and (30) shows, the notion of linking, topic continuation and given/new information are closely related.

Although demotion of the agent is an important characteristic of most passives, it cannot be the primary motivation for using the passive, as some have argued (e.g. Huddleston & Pullum 2002: 1445-1446 for short passives). One of the arguments in favour of agent demotion as the primary function is the observation that the majority of passives are agentless, with percentages ranging from 70% to 80% (e.g. Biber et al. 1999: 940, Wanner 2009: 168, Siewierska 1984: 38). The reasons for omitting the agent vary, as the following examples from Givón (2001) illustrate.

(31) He was killed in the Boer war.

(Givón 2001: 125, his 80, from Jespersen 1924: 167)

(32) The soldiers invaded the village. Soon the place was burned down.

(*ibid.* 125, his 81)

(33) The plane was brought down safely.

(*ibid.* 125, his 83, from Trout, *Venus on the Halfshelf*, 39-40)

Givón explains how in (31), the agent is “unknown or unrecoverable” (2001: 125). In (32) and (33), adding the agent is unnecessary, because in (32), it is anaphorically given and in (33), it is predictable: there is no need to explain that the plane was flown by pilots. While these are obviously correct observations, at the same time we can see that in all sentences the promotion of the patient to topic is also relevant: sentence (31) is about *he*, not the killer; (32) is about *the place*, not about soldiers; and (33) is about *the plane*, not pilots. In other words, while the reason for the exclusion of the agent is different, the effect of it is similar in each case: the topical patient is expressed as the subject. Moreover, as Siewierska (1984: 37) points out, some languages always require their passives to have an agent, which means that excluding the agent cannot be the main motivation to use a passive. In fact, she points out that in English it is also not always possible to exclude the agent phrase: compare *On his death his daughter succeeded him* and **On his death he was succeeded* (1984: 38). In conclusion, patient promotion is a defining property of all passives—long and short—while agent suppression or focus on the agent can best be seen as a consequence of this promotion.

Types of passives

There are a number of aspects to the definition of passives which are potentially relevant for the selection of passives in a corpus study; these are both related to the

structure and function of the passive, and to the nature of the agent phrase. In this section, I will briefly explain these contrasts and discuss whether these classifications are relevant or applicable to the studies in the following chapters.

One widely-used distinction is that between so-called verbal (or syntactic) passives and adjectival (or lexical) passives, with the second class generally regarded not as passives but as copular constructions (cf. Wasow 1977, Chomsky 1981). In this line of thought the two types of passives are formed differently: while verbal passives are formed in the syntax—object becomes subject—adjectival passives come directly from the lexicon—no object is selected. As Huddleston & Pullum (2002: 1431) point out, the term adjectival *passive* is misleading, but the overlap seems to be there inherently, considering that the periphrastic passive started out as a copular construction (see e.g. Givón 2006, Fischer & Van der Wurff 2006: 152-154). The distinction between the two types builds on the similarities between certain participles and adjectives, cf. examples (34)a and (34)b.

- (34) a. John was chased by the dog.
b. John was frightened by the dog.

The difference in status between the participle in a and the adjective in b can be established with the use of several tests (see e.g. Huddleston & Pullum 2002: 1436-1440, Chomsky 1981: 54-55). First, while verbal passives can only be formed with *to be*, adjectival participles can follow other verbs than *be*, as illustrated in (35)a-c.

- (35) a. John seems sad.
b. John seems frightened.
c. *John seems killed.

(Chomsky 1981: 54)

Second, adjectival participles can be used in pre-nominal position, but verbal participles cannot, cf. (36)a-c.

- (36) a. The sad man.
b. The frightened man.
c. *The killed man.

Third, adjectival participles can be preceded by the prefix *un-* (*unnoticed*), but verbal participles cannot (**unkilled*). Finally, adjectival participles can be modified by *very* or *too* (*very frightened*), in contrast to verbal participles (**too killed*).

It could be argued that only verbal passives are truly passives, in the sense that they re-arrange word order, and thus only those types of passives should be included in a study of the passive. However, the distinctions are not always straightforward. Consider, for instance, the following example, which is ambiguous between verbal and adjectival (cf. (34)b above).

(37) They were married.

(Huddleston & Pullum 2002: 1436, their 32)

Huddleston & Pullum provide the following two readings, one in which the participle is verbal—*They were married last week in London*—and one in which the participle is adjectival—*Hardly anyone knew that they were married—that they had been for over ten years*. The fact that the distinction cannot be made on the basis of the type of verb, or the form of the verb, as this example with *married* shows, provides a complication for a large-scale corpus study. In addition, for the present study, the limited number of data and the lack of native speaker intuitions means that it more difficult to make this distinction in earlier stages of the language.

The same problem exists for another functional distinction, which has been proposed to follow from the distinction between adjectival and verbal passives, namely between dynamic and stative passives. Generally speaking, adjectival passives are taken to be stative, i.e. describing a state, and verbal passives are taken to be dynamic, i.e. describing an event. Some, such as Embick (2004), make a further distinction by subdividing the category of stative into resultative and stative,

- (38) a. The metal was flattened by the smith.
b. This metal is flattened.

(Embick 2004: 365, his 25a-b)

Again, however, examples may be ambiguous. For instance, some verbal passives are stative, witness the following example from Huddleston & Pullum (2002).

(39) She is loved by everyone.

(2002: 1438, their iib).

As Huddleston & Pullum explain, *love* is a stative verb in the active and it retains this stative meaning in the passive. What makes these distinctions especially difficult is that the same form can be used in two different interpretations. In addition, although a study by Toyota (2008: 17) suggests that the passive may have changed from a mostly stative to a predominantly dynamic construction between OE and PDE, it is not clear that this functional difference is relevant to the present study. Most importantly, the choices that are relevant for information-structural purposes, i.e. choices about the order of elements (*John was killed by the cat* or *The cat killed John*) and the choice of subject, need to be made in all passives, whether they are dynamic or stative.

Just as with the question as to what constitutes a real passive, there has been a discussion about what constitutes a real agent phrase, because of a number of factors. First, although it is always a prepositional phrase with *by* that encodes the agent in PDE, it is not the case that all prepositional phrases with *by* in passives are agents, as illustrated in (40).

- (40) If the Cardinals heed Manager Gene Mauch of the Phillies, they won't be misled by the Pirates' slower start this season. [Brown Corpus]
(Birner & Ward 1998: 195, their 248b)

In fact, some *by*-phrases can be used in active clauses as well (*We talked by phone, ibid.*). In addition, passive clauses may contain prepositional phrases which do not start with *by* and which are not clearly agentive, but which can nevertheless be translated into the subject of an active, as illustrated in the following examples.

- (41) Abel was killed *out of hate*. > *Hate* killed Abel. Cause
(42) Abel was killed *with a jawbone*. > *A jawbone* killed Abel. Instrument
(Svartvik 1966: 103)

Finally, it was not until the EModE period that the *by*-phrase became the established way to express the agent (see Moessner 1994). It could again be argued that only a subclass of passives should be included in a corpus study, namely those with a *by*-phrase which represents a true agent. However, with respect to the information-structural function, there is a similar argument as for the distinction between verbal and adjectival passive: the considerations of word order and the choices that speakers make are also present in passives with a non-agentive *by*-phrase. This means that the best solution is to include all *by*-phrases that can be translated into an acceptable passive, while they do not necessarily have to be agentive¹⁵ (cf. Birner & Ward 1998: 195). In conclusion, most of the finer-grained contrasts for passives and *by*-phrases are either not directly relevant for the present study, or are too fine-grained for a diachronic corpus study.

2.3.2 The information-rearranging function of the passive

In addition to the patient promotion function of the passive, another function has been identified for long passives, namely their ability to reorder arguments to achieve a given-before-new order of information. As Siewierska (1984: 222) explains, with reference to Halliday (1968), Chafe (1970), Hutchins (1975), Hinds (1975) and Dik (1978), long passives allow for a given referent to occur as the unmarked topic or theme in a sentence, and for a new referent to occur as the information focus, an order of information which would be disrupted in the active version of the same sentence. This widely discussed function has been further substantiated for PDE with a corpus study by Birner (1996) and Birner & Ward (1998), while Seoane (2000), with a similar type of corpus study, has shown that it is already a function of the long passive in EModE.

¹⁵ This would be a reason to use the term *by*-phrase throughout, rather than the term agent phrase (cf. Huddleston & Pullum 2002: 1428). However, because of the variation in prepositions in OE and ME, I will use the term agent phrase for these periods.

Present-day English: Birner & Ward (1998)

Birner & Ward (1998), based on Birner (1996), compare the information status of the subject and the agent phrase and propose that the passive is regulated by a relative information-structural constraint: the *by*-phrase must not be more given (“more familiar”) than the subject. Their study is based on 200 *by*-phrase passives from the Brown Corpus (Kucera & Francis 1967). The following examples from Birner & Ward, one felicitous passive and one infelicitous passive, illustrate their basic claim.

- (43) a. The mayor’s present term of office expires Jan 1. *He will be succeeded by Ivan Allen Jr....*
 b. Ivan Allen Jr. will take office Jan 1. #*The mayor will be succeeded by him.*

(Birner & Ward: 1998: 200, their 253a-b)

In the first example, the subject *he* is discourse-old, referring to *the mayor* in the previous sentence. It is followed by a discourse-new *by*-phrase, *by Ivan Allen Jr.* The second example is problematic because the subject of the passive clause, *the mayor*, is discourse-new, while the information in the *by*-phrase is discourse-old, with *him* referring back to *Ivan Allen Jr.* in the preceding sentence.

Birner & Ward use a binary distinction between discourse-old and discourse-new referents, based on whether or not these referents occur in the discourse preceding the relevant sentence. Table 2.6 repeats the results of their study of the discourse-status of subject and *by*-phrase in 200 long passives. It provides evidence that three orders occur: old-before-new, new-new and old-old. There are no exceptions to the constraint; that is, no examples like (43)b above. The table also shows that the majority of the subjects and *by*-phrases in fact have a similar discourse-familiarity (37.5% d-old/d-old and 18.5% d-new/d-new).

Table 2.6 *Discourse-Status and Passivization* (Birner & Ward 1998: 201, their 4.1; percentages mine)

		Subject NP		
		D-Old	D-New	Total
<i>By</i> -phrase NP	D-Old	37.5% (75)	0.0% (0)	37.5% (75)
	D-New	44.0% (88)	18.5% (37)	62.5% (125)
	Total	81.5% (163)	18.5% (37)	100.0% (200)

Birner & Ward make two further points with respect to the function of the passive which are relevant to the corpus studies in this thesis. First, they state that it is discourse-familiarity, and not hearer-familiarity, that is relevant, meaning that the knowledge of a hearer of a referent (such as *Ivan Allen Jr.* in example (43)) does not count, but only the occurrence of a referent in the discourse. Second, Birner & Ward propose that the passive has a linking function, by having a subject which refers

back, i.e. connects, to the previous discourse (cf. the topicalization function described above). There is only one type of exception to this ‘linking function’, namely those passives which have a new subject and a new *by*-phrase, as in the following example.

- (44) An alert 10-year-old safety patrol boy was congratulated by police today for his part in obtaining a reckless driving conviction against a youthful motorist. [...] [Brown Corpus]

(Birner & Ward 1998: 204, their 257)

Sentences like these usually occur in the beginning of a discourse, where it is more likely that all information in the sentence is new: the function is to introduce the discourse, rather than continue the discourse (cf. Lambrecht’s event-reporting sentences, 1994: 124). What these examples illustrate most importantly is that there is no absolute requirement that the subject of the passive be given.

In short, Birner & Ward’s study provides corpus evidence for the function of the long passive as creating or maintaining a given-before-new order of information. Interestingly, there are no exceptions in their database. It is also relevant to note that most of their examples consist of sentences where the subject and *by*-phrase have an equal discourse-familiarity, which means that there is no relative ordering in those examples. Finer-grained distinctions in information status may be able to shed more light on this.

Early Modern English: Seoane (2000, 2006)

Seoane (2000, 2006) proves that information rearranging is also a valid characterization of the function of long passives in EModE. In the majority of examples in her study, based on a corpus spanning 1420-1710, the passive obeys the same given-before-new constraint as shown by Birner & Ward (1998). Like Birner & Ward, Seoane uses a binary distinction between given and new, but unlike Birner & Ward, she bases her two categories not only on discourse occurrence but also on what Birner & Ward call ‘hearer-familiarity’. Table 2.7 repeats her results, which show that all orders occur, but that there is a clear majority (96.4%) for the orders which also occurred in Birner & Ward study: given-new, given-given and new-new.

Table 2.7 *Information conveyed by the subject and agent in passive clauses* (adapted from Seoane 2006: 378)

		Subject NP		
		Given	New	Total
<i>By</i> -phrase NP	Given	8.1% (36)	3.6% (16)	11.8% (52)
	New	64.3% (283)	23.8% (105)	88.2% (388)
	Total	72.5% (319)	27.5% (121)	100.0% (440)

Interestingly, in contrast to Birner & Ward's study, Seoane's database contains a number of examples which do not follow the constraint: 16 examples (3.6%) have a new-given order. It is unfortunate that Seoane does not comment on these examples in her article, because they could provide an interesting insight into how the constraint may be different in EModE from how it is in PDE. Specifically, it could provide an answer to the question whether these examples are really to be considered exceptions or whether the given-before-new principle is as yet less strongly present in EModE. What she does provide as an interesting example is (45), in which the patient-subject follows rather than precedes the agent *by*-phrase. In other words, the argument-rearranging function of the passive is undone, as it were, by the change in word order.

- (45) *By hure þe dewell is putt in gret drede.*
(QM3/4_IR_SERM_ROYAL:176; Seoane 2006: 379, her 14)

Nevertheless, as Seoane argues, this example provides additional evidence for the relevance of the information-structural constraint, because the *by*-phrase is in this example the more given referent. In these cases, then, the *by*-phrase provides the link to the preceding discourse, functioning as a local anchor (see Los & Dreschler 2012).¹⁶

While Seoane (2000, 2006) uses a binary given/new distinction, in a follow-up study, Seoane (2012) compares two different ways of annotating givenness in long passives, roughly corresponding to Birner & Ward's discourse-familiarity and hearer-familiarity.¹⁷ The first approach used by Seoane measures the distance from a referent to its antecedent in the text, based primarily on work by Givón (1983, 1990) and Toole (1996); the second approach is based on Prince's (1981, 1992) classification scheme, which determines the cognitive availability of referents to the hearer based both on the text and on world knowledge. Seoane concludes that Prince's approach can describe the cognitive accessibility of referents more precisely, something which is however inherent in the approach, especially when it is compared to an approach like Givón's, which considers fewer variables.¹⁸ In comparing the two approaches, she unfortunately does not address the question whether the rearranging function of the passive builds on discourse-familiarity or hearer-familiarity. Her data do suggest, nevertheless, that something may be gained from looking at finer-grained distinctions of givenness.

¹⁶ What Seoane (2006) also provides is an analysis of other aspects said to be related to information status; she finds that weight influences occurrence of long passives (not surprising since weight and information status are generally assumed to be related), but animacy and semantic roles do not.

¹⁷ Seoane's study is based on 100 long passives from Early Modern English (Helsinki Corpus, 1500-1710), Late Modern English (ARCHER, 1650-1900), and PDE (FLOB, early 1990s).

¹⁸ Seoane does not take into account anchors (as *his* in *his money*) as discourse links, something which other scales do (for instance, Birner & Ward's discourse-old category), so this may affect the results.

In conclusion, both for EModE and for PDE there exists evidence from corpus studies that the long passive functions as an information-rearranging device, by reordering arguments to achieve a given-before-new order of information. While the constraint seems to exist in roughly the same way in EModE and PDE, no such account exists for OE, and most of the studies have used coarse-grained information-structural categories while finer-grained distinctions may provide more insight into the aspects of the constraint, especially when it needs to be compared to the information-structural functions of V2.

2.3.3 A Minimalist analysis of the passive

This section presents an analysis of the passive within a Minimalist model of syntax. The analysis presented in this section builds on the similarities between passives, unaccusatives and middles, in a way that captures their shared property of lacking an external argument which expresses the thematic role of agent. The analysis of the special passives—prepositional, recipient, ECM—will be left to Chapter 3, where I discuss the analysis for each of the passives and identify the change that occurred in ME to make these passives possible.

An important first decision with respect to the syntactic analysis of the passive is whether the external argument—the subject in the active, which is left out in the passive or expressed as an adjunct PP—is initially merged or not. In other words, does the passive in (46)a start out with the same arguments as the active, as in (46)b, with a deletion of the agent in the process of the derivation, or is the agent argument not merged in the structure, as in (46)c?

- (46) a. Sagan was beaten in the sprint on the Mall.
 b. Kittel beat Sagan in the sprint on the Mall.
 c. [e] beat Sagan in the sprint on the Mall.

In the earliest version of generative grammar, the derivation of the passive was based on a full symmetry between active and passive: Chomsky's (1957) Transformational Rewrite Rule simply took an active sentence with subject and object and rewrote it as a passive (1957: 42-43, 78-81). In Government & Binding, the standard analysis of the passive (Chomsky 1981; Jaeggli 1989, Baker, Johnson & Roberts 1989) made a similar decision of base-generating the external argument, but in a crucially different way: the passive suffix *-en* was itself analysed as an argument and it acts as the external argument, in that it is assigned the agent θ -role. The external argument in turn assigns the agent θ -role to the *by*-phrase (so technically, the real external argument), through a chain which does not trigger movement. In other words, while the derivation of the passive is crucially different from that of the active from the start, what is similar in active and passive is the generation of the external argument in both versions, the processes of assignment of θ -roles and checking of Case .

Working in a Minimalist model, Collins (2005) returns to Chomsky's original rewrite rule, in that he proposes that the positions for initial merge of the external and internal argument should be similar in the active and passive: the external argument is merged in Spec,vP and the internal argument is merged in the complement-of-V position. Although Collins initially postulates that the θ -role of the *by*-phrase is checked in Spec,vP, he ultimately proposes that this checking takes place in an additional projection on top of the vP (called VoiceP). Collins gives two arguments for his proposal to merge the external argument in the passive just as it is merged in the active. Firstly, he argues against deriving the passive in a crucially different way from the active. In Government & Binding terms, this violates Baker's (1988) Uniformity of Theta-Assignment Hypothesis (UTAH) and a similar principle holds in Minimalism: θ -role assignment is configurational, meaning that a specific θ -role is associated with a specific syntactic position (2005: 83-84). Secondly, a similar initial merge in active and passive clauses captures the observed symmetry between active and passive: selectional restrictions found in the active can be copied in the passive without additional stipulations, cf. **John frightens sincerity* and **Sincerity was frightened by John* (84), an argument which was already presented by Chomsky (1957).

There are, however, several considerations that raise questions for such a strict symmetry between active and passive. Firstly, the symmetry between passive and active is by no means complete, as we have already seen in Section 2.3.2.1, which means symmetry accounts need to explain such exceptions. Secondly, the accounts above only succeed in making the derivation similar in *some* respects: while θ -role assignment in both Baker et al.'s and Collins's analysis is similar in passive and active, the derivations are not completely similar and specific operations are necessary to derive the passive and ensure that the *by*-phrase receives the appropriate θ -role. Thirdly, with respect to the selectional requirements, it is not the case that only an identical derivation can express these similarities. In particular, we need to consider the existence of verb alternations, such as the pair in (47), and more generally, the fact that verbs can be used with different complementation patterns, as illustrated in (48).

- (47) a. He opened the door.
b. The door opened.

- (48) a. He made a cake.
b. We made him president.

Specifically, the transitive/intransitive pair in (47)a and b shows that thematic roles of specific arguments can be maintained even if no completely similar initial structure is assumed—*the door* is in both cases thematically the internal argument of *open*—while (48)a and b show that verbs can more generally be used with different argument structures; no separate operations are usually assumed for these

alternations. In configurational approaches (cf. Hale & Keyser 1993, 2002), these differences are accounted for by stating that the semantic relations are expressed based on the position in which arguments are merged. Considering the existence of such alternations, it seems unclear why the passive/active alternation is treated differently by assuming that they start out as identical structures.

An alternative which allows for a view of the passive as an independent structure instead of a derived active is provided by accounts such as Marantz (1984), Kratzer (1996) and Embick (2004). These authors propose that the external argument is not initially merged in the passive, following a long-standing tradition of analysis for unaccusatives and middles (e.g. Hale & Keyser 1993, 2002). Generally speaking, these accounts postulate that no external argument is merged in Spec,vP and that in the absence of an external argument the internal argument moves to the subject position.

The possibility for clauses to be merged without an external argument has generally been used in the literature to account for transitive/intransitive verb alternations (see Hale and Keyser 1993, 2002, Levin & Rappoport Hovav 1995 and references cited there). Levin & Rappoport Hovav (1995) build on the Unaccusative Hypothesis, first proposed by Perlmutter (1978) and later adapted by Burzio (1981), which divides the class of intransitives into two categories: unergatives and unaccusatives. Unergatives, such as *sleep*, take a single external argument, while unaccusatives, such as *come*, take a single internal argument. In both cases, this single argument subsequently moves to the surface subject position. Hale and Keyser (2002) extend this type of analysis to the analysis of other transitive/intransitive alternations, such as in (49)a and (49)b.

- (49) a. The cold turned the leaves red.
b. The leaves turned red.

Figure 2.5 Structure of transitive verb (Hale & Keyser 2002)

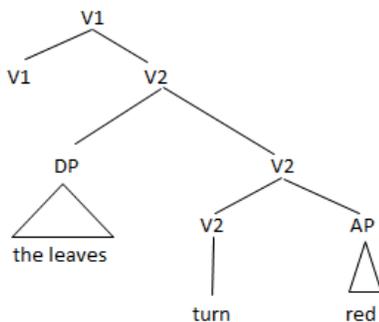
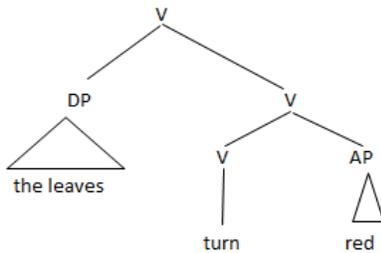


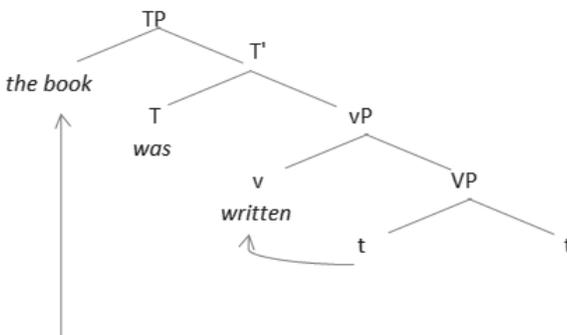
Figure 2.6 Structure of intransitive verb (Hale & Keyser 2002)



The basic structure for both verb types is similar: a verb *turn* and an Adjective Phrase *red*. In the transitive version, (49)a, the verb structure has an additional verbal layer which means it takes another specifier—the external argument *The cold*. This external argument becomes the grammatical subject of the clause; in (49)b this is the internal argument. As Hallman & Kalluli (2013: 3) point out, this configurational view of unaccusatives and transitivity alternations conforms closely to Baker’s (1988) UTAH, in that θ -roles are assigned to positions under a structural configuration.

Several authors—such as Ahn & Sailor (to appear), Embick (2004b)—have proposed a similar analysis for the passive: no external argument is merged, and as a consequence the internal argument becomes the grammatical subject. Ahn & Sailor use this type of analysis to account for middles such as *This pie bakes easily*, in which the verb also only takes an internal argument (see Chapter 6, Section 6.4.1). Building on these principles, Figure 2.7 provides the basic analysis of the passive which will be used in this thesis.

Figure 2.7 Derivation of the passive



In this analysis, the internal argument is merged in the complement-of-V position, and no external argument is merged in Spec,vP. Because the subject position needs to be filled, the internal argument moves to Spec,TP. Under standard assumptions (cf. Chomsky 1995), the lexical verb (in this case, the participle) moves to v. Note

that I assume that the auxiliary is merged in T (cf. Biberauer & Roberts 2005) and that there is no separate projection for the participle (cf. Collins who proposes a PartP as the complement of *v*, in turn taking a VP as its complement).

Having decided on this basic analysis which assumes that no external argument is merged in the passive, there are three issues that still need to be resolved: (i) how is the lack of an external argument accounted for? (ii) why does the internal argument move to subject position? and (iii) how does the *by*-phrase receive its agent θ -role if it is not related to the external argument in the active?

Until now, I have simply assumed that the external argument is not merged, but the question is how exactly to account for this alternation in structural terms. In addition, the analysis should also account for the differences between passives, middles and unaccusatives, because while they may share the lack of an external argument, they are still clearly different constructions, most noticeably in the morphology and the possibility of expressing an agent. There are broadly speaking two types of proposals in the literature which can be used to formalize a structure which lacks an external argument: using either a VoiceP or little *v* as the locus of the passive/active alternation.

The first approach is to assume a VoiceP—first proposed by Kratzer (1996)—on top of *v*P, which determines the type of clause. In some proposals this VoiceP is simply assumed and used as a landing site for elements, while no further explanation is provided for how it determines the nature of the clause. Collins (2005), for instance, makes use of such a VoiceP, but in a very basic way: his use of the VoiceP is based on evidence that there is a VoiceP in Kiswahili, as shown by Hinnebusch & Mirza (1998:111), which he assumes to be evidence for the existence of a VoiceP in Universal Grammar (UG). He then uses the VoiceP as a landing site for the passive participle and as the original merge position for *by*, but does not address the issue of how the VoiceP determines the passive character of the clause. Other approaches provide a more articulate characterization of the VoiceP and how it is involved in the formation of the passive. Ahn & Sailor (to appear) propose that the head Voice can be either [Active], [Middle] or [Passive]. The features on these different types of head determine whether a sentence is active, passive or middle. Specifically, if there is no feature for the selection of an external argument—as in the passive and middle—no external argument is merged. The problem of using VoiceP is that it assumes an additional projection, which does not, in all of these accounts, have a specific function other than stipulating the type of clause. In addition, it is not obvious that it is necessary to assume a VoiceP for active clauses. More problematically, if an active clause is said to trigger the presence of an external argument, then this account does not lend itself well to accounting for unaccusatives in a way that is similar to the derivation of passives.

A different type of approach is presented by the so-called ‘flavours-of-*v*’ approaches (cf. Folli & Harley 2005), following work, such as Chomsky (1995), that attributes a central role to *v* in determining the type of clause. Alexiadou & Anagnostopoulou (2004), summarizing earlier work, list several properties of *v*, such

as the presence of *v* in transitive and unaccusative clauses. Two crucial properties for the present analysis are first, that the features on *v* are responsible for the “licensing and interpretation of external arguments”, and second, that the introduction of an external argument by *v* is optional (2004: 119). As Alexiadou & Anagnostopoulou point out, the features on *v* combined with the presence or absence of an external argument determine the difference between a range of verb types: transitive, passives, reflexives and unaccusatives (*ibid.*). In other words, *v* determines whether an external argument is present and also determines the interpretation of this external argument when it is.¹⁹ In other proposals, these features have been specified further. For instance, Embick (2004) adds a further distinction between unaccusative and passives: he proposes a feature for a proposal of an AG (agent) feature, which is present in passives but not in unaccusatives and which gives passives the agentive “flavour” (in Hallmann & Kalluli’s 2013 words) that unaccusatives lack.

Both the so-called flavours-of-*v* (see Folli & Harley 2005) and the VoiceP approach allow for the analysis of the passive in which an external argument is not present, but the flavours-of-*v* approach presents a more insightful relation between passives, unaccusatives and middles, without the need for additional projections, instead using features.²⁰ There is one aspect that the flavours-of-*v* approach does not immediately account for, namely the morphological differences between the passive and unaccusatives. Again, this is something that can be assumed to be determined by the feature specifications on *v*. For ease of representation, I will not specify in detail the features and subheads involved, which means that in the following chapters I will abstract away from the details of the analysis. It is interesting in this respect that Embick (1998, as quoted by Alexiadou & Anagnostopoulou 2004) proposes that the difference between actives and passives is a morphological difference that does not need to be represented in the syntax.

The second question is why the internal argument moves to subject position. Most early proposals within the Government and Binding framework formalize this in terms of Case. Following Burzio’s Generalization (1986), the patient DP cannot be assigned Case within the VP (cf. also Chomsky 1981: 124), and so it needs to be assigned case in a different position. This is why the internal argument moves to the subject position (Spec,IP or SpecTP, depending on whether or not a split-IP analysis is used, cf. Pollock 1989), where it is assigned Nominative Case (Chomsky 1981, Baker, Johnson & Roberts 1989).

¹⁹ In fact, Alexiadou & Anagnostopoulou themselves propose a Voice head inside the *v*P (but crucially not a full VoiceP).

²⁰ Hallmann & Kalluli (2013) distinguish between two types of approaches: a feature approach on the one hand and a configurational approach on the other. The analysis presented here combines the two: it is partly configurational, building on the original positions of arguments, but also uses features to allow for different types of clauses. In my opinion, the two approaches do not contradict each other, but instead, the feature-approach provides a formalization of the assumptions or stipulations of a configurational approach, and in addition, allows for finer-grained distinctions in describing the type of clause.

In Minimalism, following the Probe-Goal-Agree system (cf. Section 1.3), the motivation for movement has changed from the Case-related principles in Government & Binding to the existence of an EPP-feature which triggers movement. Rather than starting from the object DP which needs to be assigned Case, it is T that starts looking (Probing) for a Goal (with matching Case and θ -features) lower down the derivation. T has an EPP feature, which essentially means that the clause needs to have a subject (cf. the original Extended Projection Principle, Chomsky 1981). The presence of an EPP-feature on T means that once an Agree relation between Probe and Goal has been established on the relevant features—i.e. valued and unvalued features on both heads which match—movement of the relevant DP to the specifier of T is triggered, movement which would not take place if the EPP feature was not present. When we consider the initial derivation of the passive as proposed here, the absence of an external argument means that the patient-DP in the complement-of-V position is the closest available argument to satisfy T's EPP, and in many cases the only argument. Compare the active and passive in Figure 2.8 and 2.9.

Figure 2.8. Derivation of active

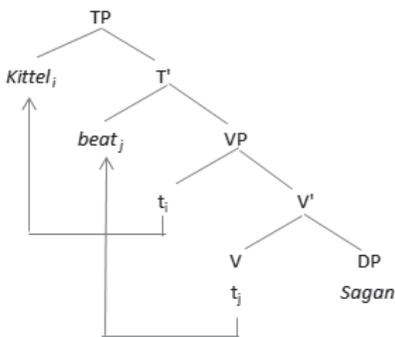
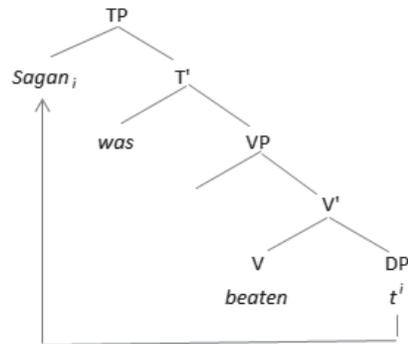


Figure 2.9. Derivation of passive



In this Minimalist configuration, Case only plays a limited role in that the object has Case features which need to be valued. This valuing can take place under an Agree relation, but it does not necessarily lead to movement. It is only because T has an EPP-feature that movement takes place (for a more detailed account of the features involved in the formation of a passive, see Platzack 2005).

The final question is how the *by*-phrase receives its agent θ -role if it is not related to the external argument in the active. In other words, if it is assumed that the agent θ -role is not assigned in the passive, then how can a *by*-phrase have this θ -role? The first thing to note in this respect is that the derivation of the *by*-phrase and its θ -role is a problem for the accounts which assume that the external argument, including the agent θ -role, are merged. For instance, in Baker, Johnson & Roberts' (1989) analysis, the θ -role is assigned to the morphology, the *-en* suffix on the past participle. They then need an additional operation to ensure that it ends up

on the *by*-phrase, which they do by stating that a chain is established between *-en* and the *by*-phrase and that by virtue of that, the θ -role is assigned to the *by*-phrase. Collins's account, which entails that the *by*-phrase is merged in Spec,vP is not without problems either, most notably that of how a prepositional phrase can be merged in Spec,vP and how it can subsequently move to its surface position in the sentence. Collins proposes that *by* is in the head of Voice, which is on top of the vP and *by* checks the case of the DP in Spec,vP (2005: 95).

Following the analysis that no external argument is merged in the initial structure and that no θ -role is assigned or needs to be assigned, I will assume that *by*-phrases are interpreted as having an agent role by a default process that interprets all prepositional phrases, in all types of clauses, in relation to the verb, including non-agentive adjuncts. The lack of an external argument in the initial structure means that other options will never be straightforward, but there are additional reasons why such an analysis is not an undesirable option. We have seen in Section 2.3.2.1 that the semantics of *by*-phrases, as well as their structural relation to the verb, are complicated, witness the fact that there are prepositional phrases that are not necessarily agents but still bear a close relation to the verb. In addition, the agent phrase is absent in the vast majority of passives, and in some cases is simply said to be an adjunct, in which case an analysis similar to adjuncts seems fitting.

In this section I have outlined an analysis of the passive which builds on the assumption that the external argument is not merged in passive clauses, something which passives have in common with unaccusatives and middles. The consequence of this is that the internal argument is the only remaining DP in the derivation which can move to Spec,TP to satisfy the EPP-feature on T—or in other words, to fulfil the requirement that every clause needs a grammatical subject. I have further proposed analysing the passive within a flavours-of-*v* approach, where the features on *v* determine the presence or absence of an external argument and possibly the agentive reading of passives which unaccusatives lack. Finally, I have proposed that the *by*-phrase in the passive—if present—behaves syntactically like an adjunct and should be represented in the derivation as such.

2.3.4 Summary and conclusions

This section has aimed to define several key properties of the passive, in order to understand both its structure and its function. What we have seen is that the passive is best described on the basis of its formal properties and most easily understood in relation to corresponding active clauses. At the same time, we have seen that there is not always a direct relation between active and passive, and that syntactically, the passive is not a derived form of the active, but similar to unaccusatives in that it lacks an external argument in the initial derivation. Most importantly, the syntactic properties of the passive—its ability to rearrange arguments from how they are ordered in the active, and its ability to place a non-agent (syntactically, an internal argument) in subject position—allow it to be used

for two specific information-structural purposes: (i) placing topical information in the preferred or unmarked position for topical information; (ii) reordering arguments so that an given-before-new order of information is achieved. These two properties make the passive a relevant alternative to V2 in performing similar functions, and a likely candidate to take over some of these functions as V2 is lost.

Section 2.4 Operationalizing information-rearranging functions

The information-structural relation between V2 and the passive as identified in the previous section builds on their shared ability to reorder arguments in the sentence, which can be used for two information-structural purposes: (i) reordering arguments to achieve an unmarked given-before-new order of information—subject and *by*-phrase in the case of the long passive, and object/PP and subject in the case of V2; and (ii) placing topical or given information in initial position, which is the unmarked and preferred position for this type of information—the subject in the passive, and the object, PP, and possibly the subject, in V2 clauses. Until now we have been assuming a working version of these given-before-new and given-first principles; this section aims to define the relevant notions in more detail. Most importantly, it aims to answer the question how givenness should be understood—as referentiality or cognitive accessibility—and how for other constructions the observations on the use of particular clause types have been translated into information-structural functions, thus providing a model for determining the function of V2 and passive throughout the history of English in the following chapters. Finally, it reviews earlier corpus studies and their annotation schemes in order to establish which information statuses are necessary to include in the corpus studies.

2.4.1 Given-before-new as the unmarked order of information

The function of reordering information to achieve an unmarked given-before-new order of information or to put the most given information in initial position can be defined in different ways, most notably based on whether givenness is purely a textual notion, or whether it should be taken to represent speaker's and hearer's knowledge. In addition, givenness is closely related to other information-structural notions, such as topic and theme, and these notions may or may not function independently of each other. This section presents a definition of the given-before-new principle as it will be used in the following chapters and addresses how the notion of markedness plays a role in the given-before-new principle.

The given-before-new principle

The observation that given information tends to precede new information is by no accounts a recent one, with early formulations going back to the beginning of the twentieth century (e.g. Behaghel 1909, 1932 and Mathesius 1928). The observations

about this order of information in the sentence have become part of the study of what is referred to as information structure, or information-packaging. An often-cited definition of information structure is that of Prince (1981), who uses the term information packaging, earlier coined by Chafe (1976):

... the crucial aspect [of given-new distinctions] seems to be the tailoring of an utterance by a sender to meet the particular assumed needs of the intended receiver. That is, information packaging in natural language reflects the sender's hypotheses about the receiver's assumptions and beliefs and strategies.

(Prince 1981: 224)

The important aspects of Prince's definition are the interaction between sender and receiver, and the fact that this interaction is inherently unbalanced (crucially different from, say, a syntactic construction), because from the beginning it is based on assumptions from the sender.

What has often been noted in these studies on information structure is a series of dichotomies: a division of the sentence into two parts, generally considered in terms of their order with respect to each other. Kruijff-Korbayová & Steedman (2003: 254) provide a detailed schematic overview of the dichotomies in the literature from the early twentieth century onwards (Russel 1905, Mathesius 1929) and how the notions used in these works—for example, topic/focus, topic/comment, theme/rheme—relate to each other. The first widely used dichotomy concerns the relations between parts of the sentence, with terms being used such as theme/rheme (e.g. Mathesius 1928, Firbas 1966, Halliday 1967) or ground/topic (Vallduví 1990). For example, theme is defined as the 'point of departure', to which information is added, the rheme (e.g. Firbas 1966, Halliday 1967, 2004). Although theme is not always defined in terms of the order of elements in the sentence, it is often the case that the theme precedes the rheme. The notions of theme/rheme are related to Lambrecht's (1994) use of the terms topic/focus, where the topic is understood as the part of the sentence or the referent about which something is added, and focus represents what is added about the topic, either new information or the establishment of a new relation (1994: 207).

The second widely used dichotomy is given/new, which considers the status of individual referents in the sentence, with the observation that given information tends to precede new information (e.g. Halliday 1967, Chafe 1976, Prince 1981, Gundel 1988). As an illustration of what this ordering principle looks like in a text, consider the following opening sentences of a newspaper article.

- (50) a. The former professional Michael Boogerd, one of the Netherlands' most popular riders, has admitted doping for a decade during his career.

- b. Boogerd says he used the banned blood booster EPO, cortisone and, late in his career, blood transfusions.
- c. The confession, in an interview with the national broadcaster NOS being aired on Wednesday, came after several reports linked the former Rabobank rider to doping practices.
(www.guardian.co.uk, 6 March 2013)

The first sentence, (50)a, by virtue of being discourse-initial, presents new information, with both the subject *The former professional Michael Boogerd, one of the Netherlands' most popular riders* and the clausal complement *doping for a decade* being new to the discourse—and considering this is an article in a British paper about a Dutch cyclist, quite plausibly new to the reader. Note that the subject, *Michael Boogerd*, has several modifiers to anchor the referent to information that is more accessible to the reader. The second sentence, (50)b, starts with *Boogerd*, who is by now given information, and continues with another clausal complement, which adds new information to the story; it further specifies the previously mentioned *doping* with new-to-the-discourse substances: *he used the banned blood booster EPO, cortisone and, late in his career, blood transfusions*. The last sentence, (50)c, again starts with old information, *the confession*, which is given in this context by virtue of being linked to the verb *admitted* in the first sentence, while both the adjunct PP *in an interview with the national broadcaster NOS being aired on Wednesday* and the complement PP *after several reports linked the former Rabobank rider to doping practices* represent new information. In all but the first sentence, then, the given information occurs clause-initially, while the new information which is added to the speaker's knowledge occurs clause-finally. This has the effect that the order in these sentences is given-before-new. Even the first sentence shows this, with the cyclist being the more accessible information and the confession that he has doped the new information (literally the 'news' in this case).

A definition of givenness

Givenness itself has been defined in different ways, partly due to the different purposes of each account which deals with givenness; for instance, some accounts use givenness to establish the relative order of elements in the discourse (Birner & Ward 1998), while others try to establish the relation between givenness and other information-structural notions, such as topicality (Lambrecht 1994), or morphological forms (Gundel, Hedberg and Zacharski 1993). Simplifying some of the contrasts, this section highlights two types of approaches to givenness, which may both be the type of givenness that characterizes the information-structural functions of the passive and V2: one that focuses on the familiarity of a referent based on text and context (e.g. Prince 1981), and another that focuses on the degree of accessibility (e.g. Gundel et al. 1993, Lambrecht 1994).

The first type of approach focuses on different ways in which referents become accessible for the hearer, based both on the discourse itself and the wider context.

Given referents in this view are understood to be known to the reader or, in other words, referents for which a mental representation exists in the reader's mind. Most of these approaches focus on a binary distinction between given or new, although many of them also recognize an additional middle category of information which is not strictly speaking given but still somehow familiar for the hearer. Prince (1981) calls this type of givenness "shared knowledge" or "assumed familiarity", referring to Kuno (1979) and Clark & Haviland (1977), and she defines it as follows: "The speaker assumes that the hearer "knows," assumes, or can infer a particular thing (but is not necessarily thinking about it)" (1981: 230). Lambrecht (1994) uses the term "identifiability", for which he gives a definition which is similar to Prince's definition for "assumed familiarity": "the speaker's assessment of whether a discourse representation of a particular referent is already stored in the hearer's mind, or not" (1994: 76). Prince points out that it is not relevant whether the information is mentioned in the discourse or whether it is only inferred from the context (1981: 231). This does not just apply to deictic references to the situation, but also to concepts or people that the speaker assumes that the reader knows. The following examples illustrate some of these categories.

- (51) a. Have you heard from **Jane-Carol** recently?
b. We got some beer out of the trunk and **it** was warm.
c. Where were **your grandparents** born?

(Prince 1981: 231, her 18a-c)

In example (51)a, the question can only be felicitous when *Jane-Carol* is known to both speaker and hearer, while (51)b is an example of a referent which is newly introduced to the discourse and which can then be referred back to in the following clause. Example (51)c, finally, shows a referent which is linked: the speaker may not know the hearer's grandparents, but the anchor *you* and the knowledge that everyone has grandparents (even though they may be dead), together make the sentence felicitous. When this type of approach is applied to corpus studies (see e.g. Prince 1981, 1992, and Birner 2006), the main goal lies in establishing the different ways in which something becomes familiar or identifiable. Section 2.4.3 explains the annotation practices in more detail.

The second type of approach to givenness is represented by models that focus on the "saliency", "predictability" or "activation" of a particular referent. These models do not make a binary distinction between given and new, but define givenness as a scalar notion and aim to capture the extent or degree to which a referent is accessible. Givenness-as-saliency is defined by Prince (1981)—with reference to Chafe (1976)—as follows: "The speaker assumes that the hearer has or could appropriately have some particular thing/entity/ . . . in his/her CONSCIOUSNESS at the time of hearing the utterance." (1981: 228). Prince adds that a referent can only be given if it has been explicitly present in the discourse before. In this approach, then, a referent cannot be given only on the basis of presence in

the hearer's mind at a random point in time. Lambrecht (1994) uses the term "activation" for a similar notion, following Chafe's (1987) definition: "the speaker's assessment of the status of the representation of an identifiable referent" (1987: 76). These methods which aim to establish the degree of givenness often use this scale to relate it to other aspects of the discourse. Lambrecht (1994), for instance, uses his notion of activation as the basis for his Topic Acceptability Scale, which expresses the degree to which a referent is acceptable as an aboutness topic in the following sentence. To give an example, a referent which is active—i.e. "in focus of consciousness"—is most acceptable as a topic because it is easy to process, while a completely new element (brand-new unanchored) requires considerable effort: a referent first has to be established before something can be said about it, otherwise the hearer cannot process the message (1994: 94). When these scalar types of approaches are applied to corpora (e.g. Gundel et al. 1993, Ariel 1988 & Givón 1983), the main goal is to establish the *degree* of givenness, based on scales of accessibility; see Section 2.4.3 for a more detailed description of the annotation practices.

We have seen in the discussion on V2 and the passive in the previous sections that both types of givenness have been used to describe (part of) their information-structural function: the distinction between subject positions seems to build on shared knowledge, as do the topicalizing function of the passive and Seoane's work on the long passive in EModE; Birner & Ward (1998), however, describe the rearranging function of the long passive in PDE specifically in terms of discourse-oldness, and the local anchoring function defined by Los (2012) and Los & Dreschler (2012) also builds on the notion of (short) discourse links. This means that in describing the overlap between the information-structural function of these clause types, it is important to consider both interpretations of givenness, but consider them separately, in order to determine the effect of each.

The relation between givenness and other information-structural notions

Given referents share properties with other information-structural notions such as theme and topic and syntactic notions such as the subject, with the result that givenness has often been equated with these notions. While the correlations are clear, however, in most cases these relations are only the effect of the notions often occurring together.

First, givenness has often been equated with theme, a term which is defined by Halliday (1967: 212) as "what comes first in the clause". With given-before-new as the general order of information in the sentence, the two notions necessarily often coincide. Nevertheless, the two concepts are not similar, as Halliday explains: "[W]hile 'given' means 'what you were talking about' (or 'what I was talking about before'), 'theme' means 'what I am talking about' (or 'what I am talking about now'); and [...] the two do not necessarily coincide." (212). In other words, where given information refers back to what has been said before, theme is forward-looking. In

addition, there exist examples which show that ‘what comes first’ does not have to be given; an example is the topicalized object in (52).

(52) FIFTY SIX HUNDRED DOLLARS we raised yesterday.²¹

(Lambrecht 1994: 295)

The initial element here, i.e. the theme, is new information. It seems, then, that there is no inherent relation between theme and givenness, but rather that they often coincide because of independent principles.

The second notion which has often been related to givenness is that of topic, specifically ‘aboutness topic’, i.e. the referent about which a statement is made in the sentence (cf. Lambrecht 1994: 131). Topic is one half of the other often-used dichotomy, topic/focus, and Lambrecht presents the generalization that the topic comes first in the sentence (199-205), and that it precedes the focus, roughly speaking defined as the information that is added about the topic referent (213). In Lambrecht’s work, topic and focus build on given (‘presupposition’) and new information (‘assertion’), respectively, but cannot be equated with them. Most importantly, a topic needs to be given information (150), but that does not mean that all given information is automatically the topic of the sentence. Other authors have also claimed that a referent must be given in order to qualify for topic status. Li & Thompson (1976), for instance, state that a topic must be definite, a notion which they describe in terms of hearer knowledge or accessibility (1976: 461). Gundel (1988: 213) also describes givenness as a condition for topichood. The correct description of the relation between topic and givenness, then, is that a topic needs to be given information but that not all given information is necessarily part of the topic of a sentence.

The final concept to overlap with givenness is the subject, because subjects are generally said to express given information. This is especially true for an SVO language like English, where the subject often coincides with theme and aboutness topic, which are often given for independent reasons. Halliday (1967: 213) indeed states that in declaratives, the subject is the unmarked theme, referring to Firbas (1966), while given-before-new is the unmarked information order (Halliday 1967: 211-212); as a result, the subject is generally given. Even in English, however, counterexamples can be found, for instance in what Lambrecht calls “event-reporting” sentences (1994: 124), such as (53).

(53) (What happened?) The CHILDREN went to SCHOOL!

(Lambrecht 1994: 121, his 4.2c)

Lambrecht explains that this sentence contains no presupposition, i.e. given information, because the function of this sentence is not to add information about a

²¹ The small caps indicate sentence accent in all examples from Lambrecht (1994).

specific referent, but to narrate something which has happened, with the consequence that all information is new. In conclusion, the relation between givenness and the notions of theme, topic and subject must be seen as resulting from independent principles which have the effect that they often—but crucially not always—coincide.

Information-structural markedness

One final issue that I would like to address here is the notion that given-before-new is the unmarked order of information. Generally speaking, this notion has been explained in two ways: (i) given-before-new is the unmarked order of information structure regardless of the syntactic status of a clause, and (ii) given-before-new is the information structure in unmarked, i.e. canonical, clauses.

The first explanation is along the lines of Halliday (1967), who defines information structure as a separate level of language and proposes that given-before-new is the unmarked order for this level. This notion of 'unmarked' is not explicitly defined by Halliday but can be understood as normal, basic or most frequent: it is the general order but exceptions are possible. In Halliday's model, then, information structure is independent of the syntax and can by itself be marked or unmarked. This view also seems to be inherent in any cross-linguistic study investigating syntactic constructions in different languages which are used to create an unmarked given-before-new order of information. One such study is Gundel (1988), who finds that the 30 languages she investigates all have means to achieve given-before-new order, but these syntactic means differ from language to language. In such a view of interaction between the information-structural level and the syntactic means available in a language, a particular information-structural function can remain the same throughout time, while syntactic means to express them may change—for independent reasons (cf. Lambrecht 1994: 29). In the case of the passive and V2, this would mean that while at one point they shared a function, changes in the syntax—such as the loss of V2, or the introduction of new passives—may have led to a situation where the balance between these syntactic options changes.

A different view on information-structural markedness is presented by Siewierska (1984: 218-219), who explains the unmarked order of information as the order that occurs in unmarked syntactic structures. This implies a more direct link between information structuring and syntax: given-before-new is unmarked by virtue of occurring in unmarked syntactic structures. It also emphasizes the notion of markedness or canonical orders in the syntax, something which is not as prominently present in the first view. Such a view of markedness requires an investigation not only of unmarked information-structural orders, but also of syntactic markedness, and it does not allow for a consideration of markedness in information structure independent from the syntax. Any change in such an interaction would be complex to conceptualize because of a two-way interaction: if a syntactic structure becomes marked, does that mean the information-structural

that is connected to it also becomes marked? Most importantly, this view seems to be mostly suitable for a synchronic study and less so for a diachronic study.

The third view is provided by Lambrecht (1994: 15-17), who positions himself somewhere between the two previous options, combining observations on unmarked word order with unmarked aspects of information structuring. With respect to syntax, he assumes that SVO is the “pragmatically unmarked” constituent order in the three languages that he discusses—English, French, Italian—while the clause-final position is the “pragmatically unmarked sentence-accent position” for lexical arguments. Because in the unmarked sequence, the preverbal subject will be topic and the postverbal object focus, the “unmarked information-structure sequence”, must be topic-focus. Unmarked in Lambrecht’s view does not mean “pragmatically neutral”—there is no such thing according to Lambrecht; neither are unmarked options more natural. Rather, they have a “greater distributional freedom”, i.e. can perform more discourse functions than alternatives and hence occur more frequently (1994: 17). Like Siewierska’s view, this builds on a direct relation between syntactic and information-structural factors, and it seems mostly suitable for describing a synchronic situation while leaving little room for change.

In the case of the passive and V2, what we have are several syntactic options (position of subject) and non-canonical clause types (passive) which are used to achieve an unmarked given-before-new order. In the case of object fronting and PP preposing in OE, we can ask ourselves whether these are syntactic options or whether they represent a non-canonical order; for PDE the answer to this question is clearer because there is a stricter canonical word order. The types of syntactic options in OE and PDE, then, are crucially different and the interaction between the options and their information-structural function must also be different. It seems that the option of seeing information structure as a separate level of linguistic description from syntax is more suitable for the present study because it allows for a stable information-structural function which can be expressed through different means as the language changes. However, this view does not allow for information-structural needs influencing syntax, an option which we also need to take into account (cf. Los 2009, 2012). Another question that needs to be considered is whether unmarked syntactic structures can have a marked information structure and unmarked information structure can occur in marked syntactic structures. From the description of V2 and the passive in this chapter, it seems to be the case that in OE, the options provided by the syntax are used to the best information-structural effect, while in PDE options for expressing information-structural needs are limited by the strict SVO order—to such an extent that the syntactic functions of subject and object themselves acquire a default information-structural status (given and new, respectively)—and other syntactic means are used to achieve the same information-structural effects.

Conclusion

In conclusion, this section has defined the given-before-new principle by discussing two different interpretations of givenness—shared knowledge (roughly following Prince 1981) and saliency/activation (Lambrecht 1994, Givón 1983)—which can both be used as a deciding factor in creating, maintaining or restoring a given-before-new order of information in the subject. Givenness in both interpretations often coincides with theme, topic and subject, but while these often occur together, there is no direct relation. In light of the comparison between passives and V2, both types of givenness seem to be relevant, while it needs to be established to what extent word order and information structure work together, and to what extent this interaction differs in OE, where word order is relatively free and more options may exist to achieve given-before-new order of information, from PDE, where the strict SVO order limits information structure. The next section turns to studies of other constructions than V2 and the passive which have been identified as having a given-before-new motivation, in order to determine the possibilities in describing the function of V2 and the passive in the following chapters.

2.4.2 Given-before-new as a function of syntactic variation

With given-before-new as the unmarked order of information, conflicts arise with the syntax in those cases where an unmarked or canonical word order stipulates an order of arguments that is not in line with a given-before-new order of information. When such conflicts arise, marked or non-canonical constructions can be used to restore the unmarked given-before-new order of information. Different types of constraints have been formulated in the literature which build on restoring or achieving given-before-new order, three of which will be discussed in this section: (i) a particular information status is a condition for movement; (ii) a non-canonical order is used to restore a relative constraint on given-before-new; and (iii) given-before-new order determines the word order where this is variable. All three applications are potentially relevant for a description of the relation between the information-rearranging functions of V2 and the passive: while the passive and V2 both obey a relative constraint, we have also seen that there might be conditions of a different type for the subject of the passive and the clause-initial constituent in V2 sentences in OE: there is the possibility that either of them need to be old or familiar, regardless of the information status of other arguments in the clause. Finally, subject placement and possibly the use of the clause-initial position show a type of variation that may be (partly) explained with reference to given-before-new but achieving given-before-new order may not necessarily be their main function.

Information status as a condition for movement

The first type of information-structural function which has been defined for the use of specific constructions is one that requires a moved element to have a certain information status. Non-canonical clause types which have been described in such a way are preposing and postposing constructions. The following examples illustrate

the two preposing constructions which Birner & Ward (1998) include in their discussion: topicalization in (54) and focus preposing in (55).²²

- (54) Colonel Bykov had delivered to Chambers in Washington six Bokhara rugs which he directed Chambers to present as gifts from him and the Soviet Government to the members of the ring who had been most co-operative.
One of these rugs Chambers delivered to Harry Dexter White. [Nixon 1962:58]
(Birner & Ward 1998: 46, their 40a; context abbreviated)

- (55) There's some fresh coffee in the kitchen. *A whole pound I made.*
(*ibid.* 85, their 97a)

The fronted elements, *One of these rugs* and *a whole pound*, are both discourse-old in Birner & Ward's annotation: *One of these rugs* is a member of the set that is mentioned in the previous sentence (*six Bokhara rugs*) and *a whole pound* (an amount) is activated by *some* in the previous sentence. What both types of preposing constructions share is that they place discourse-old information in initial position. Crucially, Birner & Ward point out that givenness when understood as hearer-familiarity—instead of discourse-familiarity—would not adequately describe the function of these constructions and that there is no condition on the relative order of information in the clause, only on the moved element.

A mirror-type function—placing new information in late, postverbal position—is identified by Birner & Ward (1998) for the postposing constructions in (56) and (57), existential *there* and presentational *there*, respectively.

- (56) "*There's a warm relationship, a great respect and trust*" between [United Air Lines]'s chairman, Stephen M. Wolf, and Sir Colin Marshall, British Air's chief executive officer, according to a person familiar with both sides.
(Birner & Ward 1998: 102, their 116a)

- (57) Famous men came—engineers, scientists, industrialists; and *eventually, in their turn, there came Jimmy the Screwsman and Napoleon Bonaparte...*
[Upfield 1950:2]
(Birner & Ward 1998: 108, their 124a)

The two postposed elements, *a warm relationship* and *Jimmy the Screwsman and Napoleon Bonaparte*, are new, or in Birner & Ward's terms unfamiliar, information. There is a difference between them, however, because while *a warm relationship* is

²² The difference between the two types has to do with the focus, as the name already indicates: in focus preposing the preposed element is focused (and accented), while in topicalization, the element that is preposed is not the focus, and both the preposed element and the focus later in the sentence are accented.

hearer-new information, Birner & Ward propose that presentational *there* depends on discourse-newness, which means that the presented information does not need to be hearer-new (although it often is). Similar to preposing constructions, then, there is a requirement on the information status of a moved element to make the use of the construction felicitous. In both cases, this condition does not seem to be connected to the syntax of the construction—both preposed and postposed elements can belong to different grammatical categories and the information-structural constraint cannot directly be translated into a syntactic requirement²³—but to the pragmatics, with information status as a consideration on a different level from the syntax.

Birner & Ward's observations illustrate two underlying aspects of the given-before-new principle, which may or may not work together: given-first and new-last. The type of constraint for the preposing and postposing constructions may be applicable to specific aspects of passives and V2, either as a condition on the element in first position as discourse-old/hearer-old (the subject of the passive or topicalized elements in V2); or as a condition that the element that occurs late in the clause should be discourse-new/hearer-new (late subjects and passive *by*-phrases). At the same time, the studies discussed in Section 2.3 suggest that at least for the long passive, givenness is not a condition on the subject of long passives because the combination of a new subject and a new *by*-phrase is also found.

Restoring given-before-new with a non-SVO order

Another type of given-before-new function related to a clause type is represented by passivization (see Section 2.3.2) and inversion; these constructions reorder two arguments according to a relative constraint which stipulates that the second element cannot be more given than the first. In the case of the long passive, this concerns the order of subject and *by*-phrase; in the case of inversion, it concerns the position of the subject and a postverbal constituent, for instance a PP, as illustrated in (58).

- (58) There are huge cartons and tins of nuts, vanilla, honey, peanut butter. Varieties of herbs tea are visible. *On the counter are loaves—whole wheat, cinnamon raisin, oatmeal, rye, soy sunflower, corn meal.* [Terkel 1974:607]
(Birner & Ward 1998: 156, their 198a).

These examples, with a PP in clause-initial position (so-called locative inversion²⁴), are the most frequent type in Birner & Ward's (1998) database, but other elements in initial position are also possible, for instance adjectives, NPs or verbs. Crucially,

²³ In split-CP models, where the CP contains several focus and topic positions, it may be possible to define this for topicalization and focus preposing, but the details of such an analysis are beyond the scope of this chapter.

²⁴ Only "so-called". The peculiarities of locative inversion will be discussed in more detail in Chapter 5.

this type of sentence involves inversion of the subject and the entire verb phrase, distinguishing it from, most notably, subject-auxiliary inversion, inversion with *there*-insertion, and quotative inversion (see Section 2.2.2 and 5.2.1).

Birner & Ward define the information-structural function for inversion—on the basis of a corpus study of 1290 tokens—in terms of discourse-familiarity: inversion “allow[s] the presentation of relatively familiar information before a comparatively unfamiliar logical subject” (1998: 165). Their notion of discourse-oldness takes into account both whether there is a reference to the previous context, and the distance to the antecedent. Table 2.1 repeats the results of their corpus study (percentages mine), which show that, similar to the passive, three of the four possible orders occur: d-old/d-old, d-old/d-new, and d-new/d-new.

Table 2.1 *Discourse-Familiarity of Preposed and Postposed Constituents* (1998: 169, their Table 4.1)

		Preposed		
		D-Old	D-New	Total
Postposed	D-Old	10.7% (138)	0.0% (0)	10.7% (138)
	D-New	78.2% (1009)	11.1% (143)	89.3% (1152)
	Total	88.9% (1147)	11.1% (143)	100.0% (1290)

The table shows that the majority of sentences show given-before-new order (78.2%), the order which most clearly illustrates the relative order prescribed by the constraint. Example (59) illustrates.

(59) They have a great big tank in the kitchen, and *in the tank are sitting all of these posts*. [Jeff Smith, *Frugal Gourmet*, 6/17/89]

(Birner & Ward 1998: 166, their 211c)

In this example, *the tank* is discourse-old information because it is mentioned in the preceding clause, and *all of these posts* is discourse-new information. Interestingly, as with long passives, there are no violations of the constraint (d-new/d-old), but what is crucially different from passives is that the category d-old/d-new is clearly the majority category, which means that in general the contrast between the two elements is larger than for the passive. This also means that the initial element is generally old (although it is not an absolute constraint), and the postverbal subject generally new.

Birner & Ward provide an additional dimension to the constraint on discourse-familiarity, by also considering saliency—operationalized as distance to the antecedent. When the initial element and the postverbal subject have the same information status, saliency licenses inversion, as illustrated in (60).

(60) Tich made tea in a blackened billy and McPherson filled a telescopic cup he took from a pocket. Seated on a form, he helped himself to sugar and then

proceeded to cut chips from a tobacco plug, the cold and empty pipe dangling from his lips against the full grey moustache. *Seated opposite him was Tich, waiting for gossip, wondering hoping.* [Upfield 1940]
(1998: 169, their 215a)

Both *him* and *Tich* are discourse-old information, but *him* is mentioned more recently than *Tich*: *Tich*'s last mention is in the first sentence, while *him* refers to *McPherson* in the first sentence and *he* in the second sentence.

Although the type of inversion discussed by Birner & Ward is syntactically different from the inversion in OE V2, it has a bearing on the late subject clauses, which may share this information-structural motivation. In addition, the functional relation between passives and Birner & Ward's inversion is evident, and this may also extend to V2 types of inversion, especially object fronting and PP preposing as used in OE. The discussion also shows that considering discourse-familiarity as a scalar notion sometimes provides additional insight into the use of a particular clause type.

Given-before-new as an explanation for word order variation

The third and final type of application of given-before-new as a function is different from the previous two because given-before-new order is not considered as the main motivation for the use of a construction, but is used to explain otherwise optional variation. The influence of information structure here seems to be more limited because it interacts with other factors, more so than in the previously described functions. One type of word order variation for which information structure has been used as an explanatory factor is the dative alternation in English: the variation of the order NP indirect object-NP direct object as in (61), and NP direct object-PP indirect object as in (62).

(61) I gave her a book.

(62) I gave a book to her.

Several studies have identified information structure as one of the determining factors, with the more given object preceding the newer object (e.g. Rappaport & Levin 2008, Krifka 2003). However, Bresnan et al. (2005), using a logistic regression model to investigate variation in the *Switchboard corpus* (Marcus, Santorini, and Marcinkiewicz 1993), provide evidence that givenness is only one of the determining factors, together with, for instance, animacy (with animate objects preceding inanimate ones). This confirms the intuition that information structure does not always by itself determine or explain variation, but interacts with other factors. This is also likely to be the case in V2 variation, and especially in the use of the passive, which has additional functions such as impersonalization and patient promotion.

Another example of an application of information status as an explanation for variation is provided by Taylor & Pintzuk (2011, 2012a, 2012b), who provide

evidence that object placement in OE subclauses is determined by the information status of the object: new objects are more likely to occur after the finite verb than given objects. Examples (63) and (64) illustrate the variation between OV and VO orders (for sake of presentation, these examples exclude one additional variable, the position of the auxiliary). The object is given in bold, the verbs are underlined.

- (63) *gif heo þæt bysmor forberan wolde*
if she that disgrace tolerate would
'if she would tolerate that disgrace'
(coalive,ÆLS_[Eugenia]:185.305; Taylor & Pintzuk 2012b: 836, their 1a)

- (64) *swa þæt heo bið forloren þam ecan life*
so that it is lost the eternal life
'so that it is lost to the eternal life'
(coalive,ÆLS_[Christmas]:144.117; Taylor & Pintzuk 2012b: 837, their 1d)

Taylor & Pintzuk (2011, 2012a, 2014) use a binary distinction between given and new, with a detailed explanation of how they categorize each of their seven subcategories (see Section 2.4.3.2). They find that 68% of the new objects occur in postverbal position, against 46% of given objects (Taylor & Pintzuk 2012a: 51, Table 2.2). In their work, they also provide evidence that this effect of information status is strongest for simple (unmodified) objects, and does not hold for complex and clausal objects: among the simple objects, 23% of the given objects occurs postverbally against 35% of the new objects (2012b: 839). For complex subjects, this is 60% against 53% (showing the opposite effect) and for clausal objects, it is 91% and 87%. In addition to recognizing complexity as a separate factor, they also find that weight is an independent factor in predicting the postverbal occurrence of the verb and does not interact with information status (2012a: 51-52). Taylor & Pintzuk's work underlines the importance of taking into account the interaction between information-structural factors and factors such as weight and length. These factors may either reinforce each other or create independent effects. It is likely that the type of given-before-new function of subject placement interacts with other factors in a similar way, i.e. that givenness is only one part of a bundle of factors that determine the position of the subject.

These two applications of the given-before-new principle are very similar to the proposals that subject placement before or after the finite verb is determined by information status (see Section 2.2.1, Bech 2001, Van Kemenade, Milićev & Baayen 2008, Van Kemenade & Westergaard 2012, Hinterhölzl & Petrova 2010). For this type of variation, it seems even more important than for the previous two functions to take into account related concepts that may have an independent influence on the word orders and may obscure effects of information status or strengthen these effects.

Conclusion

The constructions discussed here illustrate the different ways in which the given-before-new principle can be the function of a particular word order option. The type of function ranges from an absolute constraint on the one hand, as in topicalization, where the preposed element needs to be given, to relative constraints on the other hand, which can be seen both in inversion, where the fronted element needs to be more given than the subject, and in the position for the object, where given-before-new is a tendency but can be overruled by other factors. The discussion highlights the importance of two elements in the approach: (i) separating discourse-familiarity from hearer-familiarity; and (ii) the interaction between information status and factors such as weight, animacy or topichood.

2.4.3 Annotating givenness in a corpus

Over the years many annotation schemes for givenness in corpora of modern languages (and mostly English) have emerged from the numerous publications on information structure. More recently, these annotation schemes have also been used by researchers to annotate givenness in historical corpora, which present additional challenges because of the limited access to the context and the hearer's assumptions. This section outlines some notable PDE annotation schemes in detail as well as how they have been applied to OE texts, in order to determine the relevant information status distinctions for the present study.

2.4.3.1 Annotation schemes for PDE

The two definitions of givenness that exist in the literature—givenness as cognitive accessibility and givenness as occurrence in the discourse (see Section 2.4.1)—correlate with two different approaches to annotating givenness in corpora. One of the major works within the first approach, which aims to establish what speaker and hearer know on the basis of both text and context, is Prince (1981). In addition to making the previously mentioned distinction between discourse-familiarity and hearer-familiarity, she identifies three main categories of “assumed familiarity”: Evoked, Inferrable and New.

Prince's first category, Evoked, contains elements which are given or old, i.e. known to the hearer and speaker, either because of previous mention (Textually Evoked) or deictic presence (Situationally Evoked), as in (65) and (66), respectively.

(65) A guy I work with says **he** knows your sister.

(66) Pardon, would **you** have change of a quarter?

(Prince 1981: 233, her 22d and a)

In (65), *he* is Textually Evoked because the referent has been mentioned in the discourse: *a guy I work with*. In (66), *you* is Situationally Evoked because the referent is present in the speech context.

Prince's second category, *Inferable*, is used for referents which are cognitively accessible for the hearer (or so the speaker assumes), not because of previous mention, but because they are triggered by or deducible from the presence of another element, as in (67).

- (67) I got on a bus yesterday and **the driver** was drunk.
(Prince 1981: 233, her 22c)

Prince explains that it is possible to construct *the driver* from *a bus* because of the knowledge that buses have drivers. She also identifies a subset of inferables, which she calls 'containing inferables'. An example is (68).

- (68) Hey, **one of these eggs** is broken!
(Prince 1981: 233, her 22e)

In this example, the inference point for the complete referent (*these eggs*) is contained in the NP, hence the term 'containing'. Inferables are a separate category in Prince's scheme, but are, as a group, difficult to characterize along the discourse-familiarity/hearer-familiarity distinction. In fact, Birner (2006), following Haviland & Clark (1974), makes a further distinction between elaborating inferables, as in (69), and bridging inferables, as in (70), and categorizes these differently.

- (69) The house was particularly spacious. Set well back from the road, it was almost surrounded by wide lawns on which, each side of the house, grew a huge palm tree. *Beyond the right-hand palm could be seen a clothes line.*
[A.W. Upfield, *The Widows of Broom*, 1950; Charles Scribner's Sons re-print, New York, 1985, pp. 110-111]
(Birner 2006: 40, her 10b)

- (70) Mary took the picnic supplies out of the trunk. The beer was warm.
[Haviland and Clark 1974]
(*ibid.* 40, her 11)

The difference, Birner explains, lies in the direction of the reference: forward-looking elaborating inferences prepare, at the time of utterance, for pick-up in the following sentence, while backward-looking bridging inferences are only made when the element *the beer* is mentioned. Birner proposes categorizing elaborating inferences as discourse-old/hearer-old, i.e. together with referents which are mentioned in the discourse, and bridging inferences as discourse-old/hearer-new (2006: 40-42).

Prince's final category, 'New', is used for referents that are newly introduced to the discourse, either as Brand-new or Unused referents. In the case of Unused

referents, the referent is assumed to be accessible but it has not actually been mentioned in the discourse. Example (71) illustrates.

(71) **Noam Chomsky** went to Penn.

(Prince 1981: 233, her 22b)

Noam Chomsky, discourse-initial, is new to the discourse but assumed to be in the hearer's knowledge. The Brand-new referents are further subdivided into anchored, as in (72), and unanchored, as in (73).

(72) **A guy I work with** says he knows your sister

(73) **A person** bought a Toyota.

(Prince 1981: 245, her 33d and e)

While both referents are new to the discourse, in (72) the brand-new entity is linked ('anchored') to another entity in the discourse, *I*. In summary, Prince's classifications allow for three general information status distinctions: discourse-old/hearer-old, discourse-new/hearer-old, discourse-new/hearer-new, and a separate category of inferables, which contains subtypes that behave differently with respect to discourse-familiarity and hearer-familiarity.

Prince's approach specifies different ways in which a referent becomes accessible (or remains new), but what it lacks is a degree to which something can become given. Other schemes on cognitive accessibility which build on Prince's distinctions have incorporated exactly this aspect, such as Lambrecht (1994). Lambrecht's (1994: 94) identifiable referents (roughly corresponding to given) are categorized according to their 'activation' state, in turn based on Chafe's (1987) distinctions: active (currently in use), accessible (short-term memory), or inactive (long-term memory). A different type of scale which combines cognitive accessibility with a scalar approach is Gundel et al.'s (1993) Givenness Hierarchy, which aims to link cognitive states to particular referential expressions. While some of the insights from these two models may be relevant, Gundel et al.'s approach is not directly applicable to other languages because it builds on specific referential expressions in PDE, such as demonstratives and articles, and as is well known, the demonstrative system in OE is considerably different. Lambrecht's model, in turn, is more theoretical and is difficult to operationalize for purposes of annotation. Most importantly, including a degree of givenness as part of the cognitive accessibility schemes does not allow for a separate consideration of the two factors.

The second approach for annotating givenness focuses on the occurrence of referents in the discourse, and, specifically, the distance between a referent and its antecedent. One of the relevant annotation practices in this regard is Givón

(1983)²⁵, who provides a measure for referential distance, determined by counting the number of clauses between a referent and its antecedent. The value is simply given as a number: a value of 1 when the antecedent occurs in the immediately preceding clause, a value of 2 when it occurs in the clause preceding that one, etc. The maximum value is 20; this value is also assigned when no antecedent is present in the text. Although the choice for 20 clauses is somewhat arbitrary, Givón also states that short-term referencing is “the crucial psychological correlate” of referential distance and the window of relevant clauses should therefore be kept relatively small (1983: 13). The values can be used to calculate an ‘average referential distance’ for a category of elements, such as the subject or topic (for an example, see Givón 1983: 32). An adaptation of Givón’s measure is provided by Gregory & Michaelis (2001), who instead of the 20 clauses only consider 5 preceding clauses and do not assign values but use three categories: prior mention, member of a previously mentioned set, and no prior mention. These choices mean that the possibility of calculating the average referential distance is lost, but a separate annotation type for members of a previously mentioned set is a valuable addition.

While Seoane (2012) found that Givón’s approach was less useful in describing the function of the passive than Prince’s (1981, 1992) approach (see Section 2.3.2), the referential distance approach can provide an indication of the degree to which an element is given in a very straightforward way. For the present study, this insight into the discourse behaviour of referents—distinctive from accessibility—is especially relevant because of the function of local anchoring, where the distance to the antecedent is hypothesized to play a role. In addition, we have seen in the previous sections that in some functions, the discourse-familiarity and sometimes the relative discourse-familiarity—i.e. saliency—determines the use of a particular construction.

In conclusion, Prince’s (1981, 1992) scheme provides a basic categorization of different types of information status of referents in the discourse, while also addressing some of the different ways in which referents can become accessible. Givón’s (1983) scheme adds to this a degree of discourse-oldness, i.e. distinguishing between a referent which is mentioned recently, or a number of clauses preceding. The combination of cognitive accessibility and referential distance provides a good basis for the annotation of the two relevant types of givenness as identified in the previous sections.

²⁵ Givón (1983) provides measures for what he calls ‘topic continuity’, which capture the availability of referents, essentially comparable to givenness and not topic in the sense of aboutness topic. While he acknowledges that identifiability is clearly influenced by factors outside the text, what he focuses on in his 1983 study are only the textual factors, the grammatical expressions and the behaviour of referents in the discourse.

2.4.3.2 Annotation schemes for older (Germanic) languages

The measure of referential distance can be applied to other languages without problems, but the schemes to determine cognitive accessibility that were discussed in the previous section were devised for PDE and are not necessarily applicable to older stages of the language. In recent years, however, several projects have started to annotate information status in historical corpora, most notably the following:

(i) Haug et al. (2009, 2014), who annotate information structure in Greek, Latin, Gothic, Classical Armenian and old Church Slavic, in the PROIEL project ('Pragmatic resources in old Indo-European languages');

(ii) Taylor & Pintzuk (2011), who use a binary distinction of given/new for their OE data but provide a detailed explanation of subcategories, and Taylor & Pintzuk (2014), who compare different schemes that are available and test it against an information-structural diagnostic;

(iii) Los & Komen (2012) and Komen et al. (2014), who present the 'Pentaset', a system of information status primaries, which together with syntactic information provides the necessary details to study information structure in OE.

This section reviews these annotation schemes in order to determine which contrasts will be necessary to include in the corpus studies in the following chapters. Rather than discussing each scheme in detail, I will focus on comparing these schemes and highlighting the basic categories that they all distinguish, roughly based on four of Prince's (1981) original contrasts: discourse-old/hearer-old, discourse-new/hearer-old, inferables, and discourse-new/hearer-new.

First, all schemes distinguish a category for information that is mentioned in the discourse, corresponding to discourse-old/hearer-old or *Textually Evoked* in Prince's work. The categories are called *Identity* in Komen et al. (2014), *OLD* in Haug et al. (2014), and *Previously mentioned* in Taylor & Pintzuk (2011). Example (74) provides an OE example.

- (74) & of *ðæs treowes wæstmæ þe is on middan neorxnawange, God bebead us, ðæt we ne æton, ne we **ðæt treow** ne hrepodon ði læs ðe we swelton.*
 and of the tree's fruit which is in middle paradise God bid us that we NEG eat nor we the tree NEG touched lest we die
 'and of the fruit of the tree which is in the middle of Paradise, God bid us that we may not eat, nor may we touch **the tree** lest we die.'
 (cootest, Gen:3.3.123; Taylor & Pintzuk 2014: 58-9, their 4)

In (74), *ðæt treow* refers back to an earlier mention of the tree in the first part of the sentence, *ðæs treowes wæstmē*. Since we have seen in the previous discussions that discourse-familiarity may be a different consideration from cognitive accessibility, it is important to separate this type of discourse-old information from other types of accessible information. Most importantly, the proposed function for local anchoring (Los 2012, Los & Dreschler 2012) is characterized not only by encoding accessible information, but specifically by establishing links to the immediately preceding discourse. While none of these schemes specify the distance or window, this can be done using Givón’s measure for referential distance.

Second, all schemes recognize at least one category for referents that are not mentioned in the discourse but are still somehow cognitively accessible to the reader, corresponding to Prince’s discourse-new/hearer-old, which contains both *Situationally Evoked* referents and *Unused* referents. Komen et al. (2014) call this category *Assumed*; Taylor & Pintzuk (2011) use the label *shared/cultural knowledge*; and in PROIEL, there is a choice between the label ACC-SIT, used for deictic reference (the situation), or ACC-GEN, used for world knowledge (encyclopaedic knowledge). Example (75) illustrates an instance of world knowledge.

- (75) *Onð þeah þe wærgcweodole Godes rice gesittan ne mægen*
 And although those-that-curse God’s kingdom occupy NEG may
 ‘And although those that curse may not occupy **God’s kingdom** ...’
 (cobede, Bede_4:27.356.26.3595; Taylor & Pintzuk 2011: 76, their 8)

The phrase *Godes rice* does not need to have an explicit mention in the preceding discourse to become accessible to the reader: it can be assumed that someone reading *Bede* knows the concept of *God’s kingdom* from the broader cultural context.

Another group of referents that can be classed under discourse-new/hearer-old is the category of generics, defined by Taylor & Pintzuk (2014) as “reference to kinds (classes/types) or specimens of kinds, rather than to individuals or groups” (2014: 60). Taylor & Pintzuk (2011), as well as PROIEL (using the tag KIND), separate generics from the category *world knowledge*, while Komen et al. (2014) do not. Example (76) illustrates a generic referent in New Testament Greek.

- (76) *ouk ep’ artōi monōi zēsatai ho anthrōpos*
 not by bread alone will-live the man.KIND
 ‘Man shall not live by bread alone.’
 (Mt 4:4; Haug et al. 2014: 36, their 38)

The subject of this example is not one specific person, but mankind in general. Taylor & Pintzuk (2014) test the distinction between the two types of hearer-old information—discourse-old and accessible—using the diagnostic of postverbal or preverbal occurrence of objects, which in previous work (e.g. 2011) they have

shown to be sensitive to information status, with new objects more likely to show VO order (cf. Section 2.4.2). They find that there is a difference between the two types of old discussed here: discourse-old referents show 31.1% VO order, while the various types of accessible information (generic / general knowledge / situation) show 38.3% VO, with only small differences between the three categories; however, the differences are not statistically significant. These findings once again underline the importance of recognizing discourse-old referents as a separate class, while the data also indicate that the different types of accessible information do not need to be considered separately but can be taken to behave roughly as one group.

The third category that is distinguished by all schemes is that of inferables. The Pentaset and PROIEL have one category of inferables (*Inferred* for the Pentaset, ACC-INF for PROIEL). Komen et al. (2014) define inferables as referents which have a textual link, but instead of an *Identity* link, they have for example a part-whole relation, as in (77).

- (77) a. My father was **one of the three sons** of Captain J. Fayerer:
 b. **the eldest** was the Rev. Joseph Fayerer, rector of St Teath, Cornwall;
 (Komen et al. 2014: 88, their 10c,d,e)

The referent *the eldest* in (77)b is part of the set *one of the three sons* in (77)a. Taylor & Pintzuk (2011) distinguish between two types of inferables, following Birner (2006): elaborating inferables—as in (78)—and bridging inferables—as in (79).

- (78) *Healdað mine bebodu & mine domas ðæt ge libbon eower lif butan ælcere sorhge.*
 Keep my commands and my judgements so-that you may-live your life without any sorrow
 ‘Keep my commands and my judgements so you may live **your life** without any sorrow’
 (Lev:25.18.3864; Taylor & Pintzuk 2011: 76, their 6)
- (79) *ðæt is, ðæt ic sette minne renbogan on wolcnum*
 that is, that I set my rainbow in clouds
 ‘that is, that I set **my rainbow** in the clouds’
 (cootest, Gen.9.13.388; Taylor & Pintzuk 2011: 77, their 10)

Both *eower lif* and *minne renbogan* are anchored to a referent in the text, but the elaborating inference in (78) is accessible without the anchor, while the bridging inference in (79) is not; Taylor & Pintzuk therefore class elaborating inferables as given and bridging inferables as new. They point out that elaborating inferables in their database are often body parts or inalienable possession, like *your life* in (78), while bridging inferables are often alienable possession. In Taylor & Pintzuk (2014),

evidence is provided for this classification, based on their diagnostic of object placement: elaborating inferences have 30.4% VO order, comparable to d-old/h-old referents, while bridging inferables have 45.1% VO order, which is closer to d-new/h-new referents (2014: 68). Since this difference is directly reflected in their results and crosses the border between given and new in a binary system, it is necessary to separate them.

The final category recognized by all schemes is that of new referents, subdivided by Prince into anchored and unanchored referents. This subdivision is not followed by all schemes, but some schemes add other types of new referents. Komen et al. (2014), for instance, make a distinction between *Inert* and *New*: in contrast to a ‘regular’ new referent, an inert referent is not available as an antecedent for the following sentence. They give the following example.

- (80) the third, Edward, a midshipman in the navy, was drowned when H.M.S. Defence foundered, with all hands, in a gale of **wind** in the Baltic in 1811.
(Komen et al. 2014: 88, their 10e)

Komen et al. argue that *wind* in (80) cannot be picked up in the following sentence because *wind* describes the gale but is itself not an independent referent. Taylor & Pintzuk (2011) have a similar category, but one which is defined specifically for objects, called semantic incorporation; example (81) illustrates.

- (81) *þæt hi moston wite þrowian for criste*
that they might torture suffer for Christ
‘that they might suffer **torture** (i.e. ‘be tortured’) for Christ.’
(coalive,ÆLS_[Chrysanthus]:216.7456; Taylor & Pintzuk 2011: 77, their 11)

The noun *torture* constitutes a unit with the verb *suffer*, and is not an independent referent which can be referred to in following sentences. Another distinction for new referents is made by PROIEL and Taylor & Pintzuk (2014) between specific and non-specific referents. Taylor & Pintzuk show that specific referents behave like d-new/h-new referents and non-specific referents behave like d-old/h-old referents in terms of object placement with respect to the verb. They point out, however, as PROIEL do, that the distinction between non-specific—classified as new—and generic—classified as given—is difficult to make, something which may affect the validity of the annotation. Finally, Taylor & Pintzuk (2011, 2014) and PROIEL, with reference to Karttunen (1976), add a separate category for short-term referents, which can only be referred to within a limited context. Example (82) illustrates.

- (82) *Ðeah þe hwa wille her on life habban gode dagas, he ne mæg hi her findan*
Yet whoever will here in life have good days, he NEG can them here find
‘Yet whoever will have **good days** here in life, he cannot find them here.’
(coalive,ÆLS[Ash_Wed]: 82.2748; Taylor & Pintzuk 2011: 76, their 5)

Considering the categories of new referents, it seems that the distinction between anchored and unanchored is relevant for the present study because of the difference in discourse behaviour: while these referents may be new in terms of cognitive status, they are still linked to the discourse through the anchor. The remaining distinctions between types of new information will be less relevant because the distinctions are too fine-grained for the present purposes.

In conclusion, Prince's four main categories—old, accessible, inferable, new—are distinguished by all schemes, although some schemes make further distinctions, most notably between types of accessible referents and types of new referents. The following table provides an overview of the information status categories discussed in this section.

Table 2.3 Comparison of annotation schemes for OE to Prince (1981)

	Prince (1981)	Haug et al. (2014)	Komen et al. (2014)	Pintzuk & Taylor (2011, 2014)
d-old/h-old	Textually Evoked	OLD NONSPEC-OLD	Identity	Previously mentioned
d-new/h-old	Situationally Evoked Unused	ACC-SIT ACC-GEN ACC-INF KIND QUANT NONSPEC-INF	Assumed	Shared/cultural knowledge Generic Situationally Evoked
Inferable	Containing inferrable Noncontaining inferrable	ACC-INF	Inferred	Elaborating inference Bridging inference
d-new/h-new	Brand-new anchored Brand-new unanchored	NONSPEC	Inert New	New Specific Non-specific Short-term referents Semantic incorporation

Following this discussion of annotation schemes, and given the aims of the current study, three contrasts need to be included in the corpus studies in addition to a binary distinction between given and new: discourse-old vs accessible, bridging vs elaborating inferables and anchored vs unanchored referents. This selection means that there are also some contrasts that seem to be less relevant for the present study, most notably between types of accessible information and types of new information.

2.4.4 Summary and conclusions

This section aimed to define in more detail the notion of ‘information-rearranging potential’ of the passive and V2 that was identified in the previous sections and to determine the most suitable approach for annotating givenness in a(n historical) corpus. Most importantly, this section has shown that two widely-used interpretations of givenness—cognitive accessibility and saliency—both need to be considered, and that several information-rearranging functions exist which could provide the characterization of the function of the passive and V2: absolute conditions on the information status of a moved element, a relative constraint for two elements in the clause, or information status as one of the factors determining word order variation. Finally, this section addressed the annotation of givenness in a corpus, for which both PDE and OE schemes that are available in the literature were discussed. Most importantly, this section has provided the definitions and method necessary to establish the information-structural potential of passives and V2 through a corpus study of various stages in the history of English.

Section 2.5 Conclusion

This chapter has aimed to provide a more detailed explanation of the hypothesized relation between passives and V2, while at the same time reviewing the possible contrasts and functions that will be relevant to the studies in the following chapters. I have highlighted the information-structural aspects of the V2 system as it was in place in OE, as well as the information-structural functions of the passive. In addition, I have discussed several issues of operationalization that provide the basis for the study of these functions and their interaction in the following chapters.

The first part of this chapter showed that the properties and especially the variation in the OE V2 system are difficult to explain with reference to syntactic principles only and that in recent years information-structural explanations have been proposed to account for this variation: most crucially, given subjects are said to occur before the finite and new subjects after the finite verb. The V2 constraint was lost roughly in the fifteenth century, although the developments differ per context: the type of clause-initial element and type of verb are two of the important factors that need to be taken into account. I also discussed several recent proposals (Los 2009, Seoane 2006, Los & Dreschler 2012, Komen et al. 2014), which argue that the information-structural properties of the presubject position as well as that of the subject changed and that this led to a greater importance of the subject, because the subject remained as the only unmarked theme. This greater importance of the subject is in turn likely to have led to an increase in the use of specific strategies for speakers to create more of these subjects, which is where the passives, middles and non-agentive subjects come in.

The discussion of the properties of the passive showed that the passive is often considered in relation to active syntactic clauses and that a definition of the passive in syntactic terms soon leads to a consideration of its function. The most basic

function is to promote the patient, which reflects the syntactic differences between active and passive. In information-structural terms, this has been defined in two ways: (i) topicalization—placing the aboutness topic or the most given element early in the sentence—as the main function of short passives; (ii) long passives—which rearrange patient and agent—as rearrangers of the information in the clause, supported by a corpus study for PDE by Birner & Ward (1998), and for EModE by Seoane (2000, 2006). I argued against a syntactic analysis which takes the passive to be a derived active, and connected the passive as a speaker’s choice in the selection of arguments to the derivation of unaccusatives and middles.

The final section in this chapter aimed to define the given-before-new principle as the unmarked order of information and provided a principal definition of givenness as shared knowledge and saliency. I also showed how the principle of given-before-new order of information has been used to account for the use of several constructions, with either givenness or discourse presence—or in some cases both—as the main motivation for the choice for a particular construction. Finally, this section aimed to identify the relevant information status contrasts based on previous annotations of givenness: four general categories were reviewed that were used both in PDE as well as OE studies: discourse-old, hearer-old (but discourse-new), inferable, and hearer-new.

In conclusion, this chapter has identified the shared information-rearranging potential of passives and V2, while at the same time explaining how these functions may have interacted in the history of English, as well as reviewing ways in which these functions may be defined in more detail. Most crucially, one of the questions that need to be answered is whether discourse-familiarity or hearer-familiarity is most important for these two constructions in OE and later periods, while at the same time the interaction between information-structural functions and syntactic principles needs to be established. The studies in the following chapter will establish which of the functions discussed in this chapter are most applicable to the passives and V2 in the history of English, and which of the information status contrasts capture these functions most accurately.

3. The Introduction of New Passives in the Fourteenth Century

Section 3.1 Introduction

This chapter presents the results of a study of the introduction of the new passives in the fourteenth century: the prepositional passive, the recipient passive and the Exceptional Case Marking (ECM) passive. It aims to establish the time of introduction as well as the reasons for introduction of each individual passive. While these issues have been widely debated, the syntactically annotated corpora that have become available since 2000 allow for a re-examination. In this chapter, I argue that the new passives were only very marginally around before the first stages of the loss of verb second (V2), i.e. around AD 1400, and that their introduction was due to very specific circumstances in the context for each passive. This means that the introduction of these passives cannot be considered as one larger development, due to the loss of V2.

The new types of passives are illustrated in (1)-(3). The passives in the (a) examples seem to be the result of the passivization of a prepositional object in (1), the passivization of an indirect object in the recipient passive of (2) and the passivization of an ECM object in (3). The (b) examples show the corresponding active sentences.

- (1) a. The pocket is being reached for.
b. The referee reaches for his pocket.
- (2) a. She was given a bracelet for her birthday.
b. Peter gave her a bracelet for her birthday.
- (3) a. Seven hundred people are reported to have died.
b. Some report seven hundred people to have died.

These passives are both grammatical and productive patterns in English, even though their use is restricted for lexical reasons—they all depend on the availability of a specific type of verb: verbs taking prepositional objects for the prepositional passive, ditransitive verbs for the recipient passive, and ECM verbs for the ECM passive. What makes these passives stand out is not only the fact that they were introduced around roughly the same time (cf. Denison 1993, Allen 1995 and Warner 1982), but also that they represent a unique feature of English: none of its closely-related languages have similar passives, and the passives are cross-linguistically rare beyond the Germanic languages (see Denison 1993: 103, 125). These observations prompt the question of the origin of these special passives and their development throughout the history of the English language.

The prepositional passive and recipient passive feature in the earliest syntactic works on historical syntax, such as the detailed studies by Van der Gaaf (1929, 1930) and Visser (1973), who both claim that these passives were both introduced in the thirteenth century. Detailed scrutiny of the early examples has led Denison (1985, 1993) to conclude that the first example of the prepositional passive dates from the early thirteenth century, while Allen (1995) concludes that the first example of the recipient passive dates not from the thirteenth century but from the late fourteenth century. The introduction of the ECM passive, on the other hand, represents a different type of scenario: it has mostly been considered in relation to its active counterpart, which itself represented an innovation and was introduced in English in the late fourteenth century (Warner 1982, Fischer 1989).

The exceptional nature of these passives, in combination with their respective times of introduction—which seem to lie within the span of one century—raises the question whether the use of the new passives represent a single development. In other words, the time of introduction and the structural similarities might suggest that a change took place in the passive construction during this period, in that the passive could now target not only direct objects, but also prepositional objects, indirect objects and lower-clause objects. Such a scenario has been proposed for recipient and prepositional passives by Lightfoot (1981) and Van Kemenade (1987) in terms of the loss of inherent case. Case, however, has come to play a less important role in the passive in generative syntax with the introduction of the Minimalist Program. The necessary first step towards answering the question whether the new passives represent one development is to re-assess the structural

connection between the passives, and to reconsider how close together the times of introduction in fact are.

The timing of the introduction of the passives also needs to be considered in relation to the timing of the loss of V2, following the proposed connection between the loss of V2 and the introduction of these passives as outlined in the previous chapters. The scenario proposed by Los (2009) sees the need for new subject strategies, and crucially unmarked subjects, result in the introduction and/or establishment of new passives. Timing alone, however, is not a strong enough argument to prove this connection; in investigating the relation between the loss of V2 and these new passives, we also need to consider the possibility of a structural relation. While Chapter 4 and 5 study in more detail the changes in the Left Periphery said to have influenced the subject and indirectly the role of the passive, the relevant part for the present chapter is to consider whether causes said to have influenced the introduction of the new passives are in fact aspects of the loss of V2. As an example, Allen (1995) has argued that the loss of the dative-fronted passive—a precursor of the recipient passive—is a direct consequence of the limitations of (indirect) object fronting, in its turn part of the series of events that constitute the loss of V2: the increasing limitations on fronting of datives meant that passives in which a dative object is fronted were also gradually lost.

In this chapter, I argue that (i) the times of introduction of the three passives coincide more closely than previously suggested, locating the crucial changes in the second half of the fourteenth century; (ii) despite this, there is no structural innovation that ‘suddenly’ allowed passivization of different types of objects across the board; and (iii) the introduction of the new passives can be ascribed to specific changes in the context of the relevant verbs, both in the active and passive. Sections 3.2 and 3.3 present new evidence for two of the previously proposed scenarios for the introduction of the prepositional and recipient passives: the reanalysis of verb and preposition as one unit for the prepositional passive; and the loss of the dative-fronted passive to pave the way for the recipient passive. Section 3.4 provides further evidence, contrary to some proposals in the literature, that the passive ECM was introduced at the same time as the active ECM, and I propose that the passive version caught on more easily than the active version because of similar passives that were already available in English. Section 3.5 concludes.

Section 3.2 The introduction of the prepositional passive

This section investigates the introduction of the prepositional passive by re-examining the early examples given in the literature and complementing these with a corpus study of the prepositional passive in the PPCME2. I argue that the introduction of the prepositional passive must be located in the fourteenth century rather than the thirteenth century, as previously proposed in the literature. I also provide evidence for a scenario which takes the introduction of this passive to be due to a series of minimal alterations in which reanalysis—i.e. the (perceived) unity

of verb and preposition—plays a crucial role from the very first examples onwards, both from a structural as well as a language-user perspective.

The prepositional passive, for example *The pocket is being reached for*, is generally said to have been introduced in the early thirteenth century, with the earliest examples dating from 1230 (Visser 1973, Denison 1993). However, the examples from the thirteenth century are scarce and not all of them are easy to interpret. It is only from the late fourteenth-century onwards that more numerous and less ambiguous examples are attested. The problems with interpreting the thirteenth-century examples raise the question whether these examples do in fact represent evidence of a fully grammatical prepositional passive at this period, or whether, as is often the case with these early examples, they are wrongly analysed or given the wrong date. If this is indeed the case for some of these early examples, it will mean that introduction of the prepositional passive proper possibly took place later than previously assumed.

The reasons for the introduction of the prepositional passive have been widely debated in the literature (see Denison 1993 for an overview). Some of these accounts do not focus specifically on the prepositional passive, but focus on general developments which possibly influence the use of the passive, like the loss of *man* (mentioned by Van der Gaaf 1930, and seemingly supported by Denison 1993), the increasingly fixed SVX order (Denison 1985) and the loss of inherent case (Lightfoot 1981, Van Kemenade 1987). Other accounts, dealing specifically with the prepositional passive, focus on the order of preposition and verb (Van der Gaaf 1930, Thornburg 1985, Mustanoja 1960: 113, 441), or the type of verbs first used with the prepositional passive (see Denison 1985 for an account in terms of lexical diffusion). Some of these accounts seem ad hoc and there still is no clear answer to the question of why the prepositional passive was introduced in exactly this period and in the contexts in which it is first found.

One persistent issue, both in the syntactic analysis of the prepositional passive and in the use of the prepositional passive in Present-day English (PDE), is the notion of reanalysis (e.g. Chomsky 1981, Van Riemsdijk 1978, Van Kemenade 1987). This notion is based on the idea that the verb and preposition at one point come to form one unit. It is this unit, rather than only the verb, which takes an object, now a regular NP, and not—as seems to be the case—a PP. This idea of reanalysis has been central to many formal analyses in describing the mechanics of the formation of prepositional passives. It has, to my knowledge, not been used as an explanation for the introduction, despite evidence that it is already relevant from the earliest examples onwards, with some of the early examples reported by Visser (1973) and Denison (1993) showing conjunctions of a single verb with a verb+preposition unit. These signs of reanalysis can be more crucial to the introduction of the prepositional passive than previously assumed.

This section is structured as follows. Section 3.2.1 defines the prepositional passive and discusses its use in PDE, with Section 3.2.1.2 presenting a syntactic analysis of the prepositional passive, building on the mechanism of reanalysis of verb and preposition as one unit. Section 3.2.2 discusses the earliest examples of

the prepositional passive—I discount the thirteenth-century examples and give additional fourteenth and fifteenth-century examples, expanding the database of verb+preposition combinations that exists in the literature. Section 3.2.3 provides a new scenario for the introduction of the prepositional passive: I provide evidence that reanalysis of verb+preposition as one unit takes place from the earliest examples onwards, and argue that the prepositional passive became possible after first occurring in relative clauses, where it represents a minimal alteration and may not have been noticed by speakers as a grammatically different construction. Section 3.2.4 concludes.

3.2.1 A characterization of the prepositional passive

The prepositional passive is a productive construction in PDE but its use is subject to several constraints, some of which are not yet fully understood. A clear factor which determines its occurrence is a close relation between verb and preposition, for instance in terms of selection of the preposition by the verb (Huddleston & Pullum 2002: 1433); this notion does not, however, account for the grammaticality of all examples. More problematically, the perceived or structural unity of the verb and preposition cannot be straightforwardly accounted for in syntactic theory.

3.2.1.1 The prepositional passive in Present-day English

The prepositional passive is the passive of an active verb which is followed by a prepositional phrase, either an argument or an adjunct. Example (4)a illustrates a prepositional passive from a newspaper, with a reconstructed active counterpart in (4)b. In this and following examples of prepositional passives, the subject is underlined and the participle and preposition are given in bold. In the active sentences the verb and preposition are also given in bold and the NP complement of the preposition is underlined.

- (4) a. ... but the Keens have been renovating their main home and it has not been **lived in** for up to a year.
(www.guardian.co.uk, 10 July 2009)
- b. They have not **lived in** their main home for up to a year.

The complement of the preposition *in* in the active sentence in (4)b, *their main home*, has been passivized in (4)a, leaving the preposition *in* stranded after the passive participle *lived*. The usual analysis for the active examples is to take *in their main home* to be the prepositional argument of the verb, but analyses on the exact connection between verb and preposition vary—in some instances a case is made for the acceptability of a prepositional passive being limited to those instances where the verb is said to lexically select the preposition (cf. Chomsky's 1986 L[exical]-marking or Van Riemsdijk's (1978) subcategorization). It is not the case, however, that only argument prepositional phrases can be turned into prepositional passives. Huddleston & Pullum (2002:1433-4) identify two types of prepositional

passives—one in which the verb selects (“specifies”) the preposition, as in (5)a and another for which they claim the preposition is “less constrained”, such as (5)b.

- (5) a. My mother approved of the plan. The plan was approved of by my mother.
b. Someone has slept in this bed. This bed has been slept in.

(Huddleston & Pullum 2002: 1433, their 20i a-b and iv a-b)

While the preposition *of* is selected by the verb in the first example, Huddleston & Pullum state that in the second example, the preposition *in* is not selected by the verb because other prepositions can also be used in this sentence: *Someone has slept on/under/near the bed (ibid.)*. Phrasal-prepositional verbs sometimes also allow prepositional passive (*The main goals seems to have been lost sight of*, Huddleston & Pullum 2002: 1433, their 20iib), but I will focus on sentences with only a prepositional object.

Most accounts which attempt to explain the acceptability or grammaticality of prepositional passives by pointing to the connection between verb and preposition state that the two elements must constitute some kind of unit (Van Riemsdijk 1978, Hornstein & Weinberg 1981). This idea of unity is further supported by the existence of fixed collocations, with verbs like *live* often occurring with the same preposition, *in*. Van Riemsdijk (1978) provides three suggestions for formalizing the notion of ‘semantic units’. The first suggestion is the notion of ‘subcategorization’, which entails that the head (the verb in this case) stipulates what type of complement it occurs with (what type of preposition in this case). To illustrate how this explains the use of prepositional passives, Van Riemsdijk contrasts (6)a with (6)b, where the second example ‘subcategorizes’ for the preposition less closely than the first.

- (6) a. She was **provided for** quite adequately.
b. *His mother was **travelled with** by John.

(Van Riemsdijk 1978: 218, his 7a; and 220, his 12)

Van Riemsdijk’s example is perhaps unfortunate because *travel* cannot passivize under any circumstance, making the sentence in (2b) ungrammatical for reasons other than the lack of subcategorization by the verb. The comparison to another stranding environment, however, shows that *travel with* does allow stranding: *Who did John’s mother travel with?* is perfectly acceptable (1978: 145, his 30). Subcategorization, then, provides a more formal approach to the connection between verb and preposition and can account for some perhaps unexpected grammatical examples, such as the active-passive pair in (7), and the selection of the adjunct in (8).

- (7) a. John **dashed into** the building.
b. The building was **dashed into** by John. (Takami 1992: 103, his 44a-b)

(8) This desk should not be **written on**.

(Takami 1992: 103, his 46a)

In (8), the adjunct *on this desk* is clearly selected ('subcategorized for') by the verb *write*, as one of the prepositions that can occur logically with *write*. At the same time, subcategorization seems to be mostly a stipulation, based on observations about frequently occurring collocations or logical options, and does not provide a further explanation as to the why of these combinations.

Van Riemsdijk's second suggestion is that the verb-preposition combinations should have a corresponding single word with the same meaning in order to allow the prepositional passive (1978: 221). Hornstein & Weinburg (1981) follow up on this idea and provide the following examples: "Thus, *talk about*, *take advantage of*, and *keep tabs on* mean 'discuss', 'exploit', and 'watch (closely)', respectively." (1981: 65). However, it is clear that this notion is less successful than the notion of subcategorization in that it cannot explain all examples of grammatical prepositional passives, witness (9).

(9) His speech was not **listened to**.

(Takami 1992: 100, his 31b)

While *listen to* may be a fixed collocation, it is not clear what single-word alternative would be available for this sentence.

The final suggestion by Van Riemsdijk, echoed by other researchers (Chomsky 1981²⁶, Goh 2000, 2001), provides a structural formalization of the connection between verb and preposition in proposing that they are reanalysed as one structural unit. Goh (2000) gives the following sentences as evidence for this reanalysis in the passive (referring to Jespersen (1909-49: part III, vol. I. 15.7)) by showing that nothing can occur between the verb and preposition in the prepositional passive, witness examples (10)c and (11)c. The (a) and (b) examples show that it is possible for material to occur between verb and preposition in the active.

- (10) a. All the committee members **agree** unanimously **on** the proposal.
 b. **On** which proposal did all the committee members **agree** unanimously?
 c. The proposal **was agreed** (*unanimously) **on** by all the committee members.
- (11) a. The search committee **asked** politely **for** additional information.
 b. This is the additional information **for** which the search committee **asked**.
 c. Additional information **was asked** (*politely) **for** by the search committee.

(Goh 2000: 156, his 3 and 4)

²⁶ Van Riemsdijk in fact refers to Chomsky (but no specific publication) for the original idea.

The fact that nothing can occur between *agreed* and *on*, and *asked* and *for* in the passive sentences suggests that their relation is tighter than in the active sentences, where it is in fact possible to split the verb and preposition.

Despite the widespread idea of the idiomatic nature of the prepositional passive and the idea of a ‘semantic unit’, however formalized, there are quite a number of examples of prepositional passives that might not be considered instances of semantic units. Why, for instance, is (a) grammatical but (b) not?

(12) The bed was **slept in**.

(13) *New York was **slept in**.

Several authors attribute the grammaticality of examples such as (a) to pragmatic constraints. For instance, Takami (1992) refers to Bolinger (1975) for a first account in terms of ‘affectedness’—not specifically for prepositional passives but for passives in general.

(14) a. I was approached by the stranger.

b. *I was approached by the train.

(Bolinger 1975: 68, as qtd. by Takami 1992: 110, his 68a-b)

While (14)a shows the affectedness of the subject, according to Bolinger, in (14)b there is only a spatial relationship between the subject and the train. He distinguishes in this manner “true patients” from subjects that “are merely located with reference to” the other arguments (1975: 68, as qtd. by Takami 1992: 110). Another indication discussed by Bolinger for the affectedness is the size of the ‘affecter’, witness the contrast in (15)a and b.

(15) a. *He was **crawled on** by a bug.

b. He was **stepped on** by an elephant.

The crawling of the bug hardly affects the subject; the stepping on by elephants is clearly a different case. Takami adds a Characterization Condition (1992: 126) to Bolinger’s observations, which states that the sentence needs to provide information about the subject to such an extent that it functions as a characterization of the subject. He uses this to explain the difference between (16)a and (16)b.

(16) a. *The stone was **stumbled over** by John.

b. The stone will be **stumbled over** if it’s not moved.

(Takami 1992: 125, his 125b and 126c)

While the first sentence does not give any information about the stone, the second sentence does, and this explains the difference in grammaticality (see also Kuno & Takami 2004: 148-153).

A second type of pragmatic constraint, related to the affectedness account, is presented by Huddleston & Pullum (2002), who state that a prepositional passive is possible only in those cases where there is a “significant property or change in a significant property of the subject-referent” (1446). This, they explain, accounts for the difference in acceptability between (17)a and (17)b.

- (17) a. The valley could be **marched through** in less than two hours.
 b. *The village hall could be **met in**.

(Huddleston & Pullum 2002: 1446, their 65ib and 65iib)

(17)a is grammatical in Huddleston & Pullum’s line of reasoning because the fact of the two hours means “something significant about [the valley’s] length and the terrain”, while (17)b adds no significant new information about village halls (2002: 1446-7). They compare this to another example, (18), where the subject is, in Bolinger’s terms, affected.

- (18) This bed has been **slept in**.

(Huddleston & Pullum 2002: 1447)

The affectedness, according to Huddleston & Pullum, lies in the fact that the bed “becomes rumpled and the sheets are likely to be treated as no longer clean enough for someone else to use them” (2002: 1447). While affectedness is perhaps difficult to formalize in clear terms, these explanations of the grammaticality in terms of affectedness show that the use of prepositional passives is not only a syntactic affair, but is determined by semantic factors as well. In short, lexical selection and the idea of syntactic and semantic units can account for a large number of prepositional passives, while in some cases additional pragmatic constraints are needed to account for the acceptability of examples.

3.2.1.2 A syntactic analysis of the prepositional passive

The discussion in the previous section has shown that there are different grammatical and pragmatic constraints that determine the acceptability of the prepositional passive. The question a syntactic analysis needs to answer is whether a similar type of object is targeted in the different contexts: a prepositional object or a direct object. The first type of analysis (e.g. Lightfoot 1979 and Chomsky 1981: 123) implies that the passive can target other objects than direct object. The second, more commonly adopted view (e.g. Van Riemsdijk 1978, Van Kemenade 1987), entails that the verb and preposition form one verbal unit taking an NP complement.

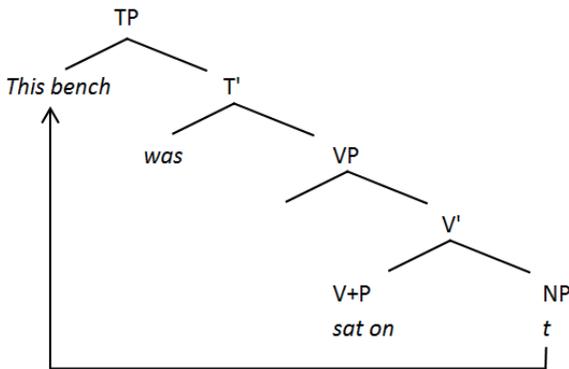
The basic idea of reanalysis for prepositional passives—as proposed by Van Riemsdijk (1978: 222) as an application of Chomsky’s (1974) proposal for a general mechanism of reanalysis—is that the verb and preposition form a larger, verbal, unit. The NP which is originally the complement of the preposition becomes the complement of the V+P unit. Example (19) illustrates the two analyses: in (a), the

verb *sat* is followed by a prepositional phrase *this bench*; in (b), reanalysis has taken place and the preposition now forms a unit with *sat*, leaving an NP *this bench*. (19)c shows the passive that is now possible.

- (19) a. Someone [_{VP} sat [_{PP} on this bench]]
 b. Someone [_{VP} [sat on] [this bench]].
 c. This bench was [_{VP} sat on] [t].

Reanalysis yields the following derivation in the passive, following the general analysis of the passive proposed in Section 2.3.3.

Figure 3.1 The prepositional passive



The verb and preposition are reanalysed as one unit and occur in the head position of the VP, which means that *this bench*, originally the complement of the preposition *on*, is now the complement of the V+P unit *sat on*. Following the analysis of the passive presented in Section 2.3.3, no external argument is merged in Spec,vP, and *this bench* is the only available candidate to move to Spec,TP, resulting in the sentence *This bench was sat on*.

The reanalysis account faces an important problem because it is not clear when and why it takes place; more specifically, whether it also takes place in the active. Van Riemsdijk's original proposal builds on it taking place in the active as well. He proposes that reanalysis of verb and preposition should be seen as a rule at the level of the lexicon: the verb and the preposition are "listed in the lexicon as semantic units" with a new label of V (1978: 222). This means that they are generated in the derivation as V[sat on]. Van Riemsdijk's proposal for positioning the reanalysis at the lexical level has the consequence that the reanalysis should also take place in the active; but it is not clear whether enough evidence exists to support this. Alternatively, they are stored as two separate items in the lexicon: the verbal unit *sit on*, and the verb *sit* (which can combine with several prepositions).

Other authors argue that reanalysis in the active is problematic. Newmeyer (2007), for instance, presents the following examples, from Jones (1987) and Inada

(1981), respectively, which he claims are problematic for a reanalysis account of stranding in *wh*-clauses, which, like prepositional passives, allow for preposition stranding.

- (20) a. Which shoes did you [walk around Europe in] ___?
 b. Which of the two knives did you [pay twice for]___?
 (Newmeyer 2007: 227, his 6a-b)

The units that the reanalysis creates, *walk around Europe in* and *pay twice for* are “utterly implausible” as lexical items, according to Newmeyer (2007: 227). He also gives the following two examples, which are passive, but in addition show *wh*-movement.

- (21) a. Which problems has Harry been [[talked to] *e* about] *e*?
 b. Who would you like to be [[sung to] *e* by] *e*?
 (Newmeyer 2007: 227, his 10a-b)

Newmeyer shows that the two reanalyses given in (21)a-b, one for *wh*-movement and one for the passive, are incompatible: they cannot be both applied in these examples. Another example from Inada (1981; as quoted by Lee 2007) also seems to provide evidence that reanalysis cannot in fact take place on a lexical level.

- (22) The bale that the goods were *packed in* went bad.

The original form of the first of part this sentence would be *Someone packed the goods in the bale*, with *packed* and *in* not adjacent, which means that they cannot have been a unit at the point of base-generation in the active. Only after *the goods* has been passivized can the unit *packed+in* be formed. While this is indeed a problematic example, the problem seems to be due to the interaction between the two types of stranding—*wh*-movement and passivization—which cannot be taken as sufficient evidence that reanalysis never takes place in the active.

Despite the various problems with reanalysis raised by these authors, there is in fact some evidence that reanalysis also takes place in the active. Interestingly, this evidence specifically addresses the OE and ME situation. Denison (1993) provides OE and early ME examples which show conjunction of single verbs with V+P collocations in the active, in turn suggesting that reanalysis has taken place, i.e. they behave and are treated as units.

- (23) & *eac symble wiðstodan & ongen fuhton heora agnum synlustum*
 and also always withstood and against fought their own sinful-desires
 ‘and also always withstood and fought against their own sinful desires’
 (GD 232.1; Denison 1993: 128, his 30)

- (24) *saul ðet is ðe ueond hateð & hunteð efter hire*
 Saul that is the devil hates and hunts for[=pursues] her
 a1250 ((*Ancr.* (Nero) 57.4); Denison 1993: 129, his 38))

In (23), the single verb *wiðstodan* is conjoined with the verb+preposition combination *ongen fuhton*, and they share the same object, *heora agnum synlustum*. Note that *ongen* is possibly paralleled with the prefix *wið* in the first verb. Example (24) shows the same structure, with the single verb *hateð* conjoined with the verb and preposition *hunteð efter* and sharing the object *hire*. These examples show that reanalysis is also possible in active sentences, although it can perhaps only be argued to have taken place if the evidence, such as conjunction, is there.

Taking into account the evidence from the active sentences in ME, I will assume that reanalysis also takes place in the active, but only optionally (indeed, following the spirit of Chomsky (1974), who, according to Van Riemsdijk (1978: 222) also presents reanalysis as *optional*). I will, however, take it to be a syntactic operation, rather than one that takes place at the lexical level. This structural analysis provides the syntactic machinery to analyse the change that took place in ME with the introduction of these new prepositional passives: we now know that we need to account for the reanalysis of verb and preposition and not for the possibility that passives target prepositional objects. But before we turn to that question, let us first consider the ME data.

3.2.2 The prepositional passive in ME

This section presents a study of the prepositional passive in the PPCME2 (Kroch & Taylor 2000) which sheds new light on the reason for the introduction of the prepositional passive. A re-examination of the early examples presented in the literature—most notably Van der Gaaf (1929), Visser (1973: 2122-2124; §§1950-1952) and Denison (1993)—in combination with the examples from the PPCME2 leads me to conclude that the introduction of the prepositional passive is later than previously proposed in the literature: in the fourteenth century instead of the thirteenth century.

3.2.2.1 Method and selection of examples

The following queries were used to search for prepositional passives in the PPCME2 corpus. As explained in Section 1.3, the searches were carried out with the programme *CorpusStudio* (Komen 2011), which makes use of the original search programme that comes with the Penn-Helsinki-York corpora, *CorpusSearch* (Randall 2005-2007). Throughout this thesis I will give the queries in the format as shown below for Query 3.1: this format abstracts away from some of the details, but it shows the order of the queries and the relevant labels and commands.

Query 3.1 Prepositional passives
a. IP* idoms *SBJ*

- b. IP* idoms *BED*|*BEP*|BE|BAG|BEN
- c. IP* idoms *VAN*
- d. IP* idoms PP* & PP* idoms P & PP* idoms NP* & NP* idoms **

Query 3.1 selects all main and subordinate clauses of all types which have a form of *to be* (*BED*|*BEP*|BE|BAG|BEN) and a passive participle (VAN) as clause constituents (line 3.1a-c). The clause also needs to dominate a prepositional phrase with a deleted NP-complement of the preposition (line 3.1b).²⁷ The query does not specify the type of trace, which means that the results include sentences with relativized or fronted elements. All examples were analysed by hand, and instances which did not represent prepositional passives were excluded from the selection. In addition, one of the results is in fact an example of a mistake in the annotation:

- (25) *Ich am made atte as a mere wyþ-outen resoun*
 I am made [atte] as an animal without reason
 1350 (cmearlps, 88.3844)

Atte is labeled as a preposition, with a following empty node as its complement within the PP. If it really is a preposition, the translation would be ‘I was made out’ and this example would represent a very early example of a prepositional passive. However, given the text of the Latin Vulgate in (26), a version of which is the source for the Early Psalter, and the King James translation in (27), it seems clear that *atte* is in fact meant to represent *at thee* rather than an obscure preposition.

- (26) *ut jumentum factus sum apud te, et ego semper tecum*
 as beast made I-am before thee, and I (am) always with-thee.
 (Clementine Vulgate, Psalms 72:23)

- (27) So foolish was I, and ignorant: I was *as* a beast before thee.
 (King James Version, Psalms 73:22)

The remaining 44 relevant examples in the PPCME2 are discussed below, together with the examples previously presented in the literature.

3.2.2.2 Earliest examples

The earliest examples according to the literature date from the thirteenth century²⁸, but they are not without problems. In fact, we will see below that the first

²⁷ In order to let CorpusStudio/CorpusSearch select traces, it needs to be stipulated in the query that traces are not ignored (i.e. the label “**” should be excluded in the categories that should be ignored “add_to_ignore”).

²⁸ There is one proposal that seems to make a different claim: Goh (2000, 2001) claims that the prepositional passive was not an innovation in the ME period but that it builds on an available structure in OE, which he calls ‘a prepositional passive in disguise’. Because Goh’s account does not actually provide

unambiguous examples and more numerous examples only start to appear in the fourteenth century.

Thirteenth century

There are three thirteenth-century examples that are candidates for the earliest attestation of a prepositional passive. The first is given by Visser (1973), from two different manuscripts, (28) and (29), as an apparent counterexample to the existing claim in the literature at his time of writing that there are no instances of prepositional passives before the fourteenth century.

(28) *heo schal beo greattre ibollen. leafdiluker leoten of þen a leafdi of hames, ...*
 she shall be more-greatly honoured, more.ladylike thought of than a lady of homes
 c1225 (Ancrene Wisse (ed. Tolkien) 58, 7)

(29) *Heo schal beo ... leafdiluker leoten of ...*
 she shall be ... more.ladylike thought of...
 ‘she shall be more greatly honoured, thought of as more ladylike than a housewife’
 c1225 (Ancr. R. (Corp-C) 58)
 (Visser 1973: 2122, §1950; translation Van Kemenade 1987: 209)

However, this example is discounted by Denison (1981, 1993), because he claims that it is not, in fact, a prepositional passive, or indeed a passive at all. First, he presents a different translation: ‘and she shall be more self-important, shall behave in a more ladylike way, than a lady of property’ (1993: 142, for his 86²⁹), although still allowing for *think of* as the right translation—indeed, the MED includes this example under the meaning *think of* for *leoten of*. A more important argument by Denison is that he questions both the status of *leoten* and *of*, proposing that *leoten* is an infinitive and *of* is elliptical, meaning that it probably did not have an object to begin with, or that the status of the preposition-like element as a preposition is not evident (1993: 161, note 11). Denison does not comment on the reasons why *leoten* would be an infinitive rather than a past participle combining with *beo* in the

earlier examples of the prepositional passive, but only a precursor, I have not included it here in the discussion. Besides, the details of his account are not convincing. Most importantly, at least a substantial part of the examples of what Goh claims to be a ‘prepositional passive in disguise’ in OE do not seem to constitute passives at all. Consider, as an illustration, the following example, which is a perfect rather than a passive.

hie þeah swa undrædendlice gebidon þæt se ege ofer-gongen wæs,
 they however fearfully awaited that the terror over-gone was
 ‘however, they fearfully awaited for the terror to be passed over’
 (Or 160.30-1; Goh 2001: 212, his 19)

²⁹ For support, Denison points to a French translation, which has ‘behave more like a lady’ (1981: 75, fn 63; 1993: 161, end note 11) and another instance in the *Ancrene* (1993: 161, note 11, *Ancr.* 36a.18).

previous clause, but does provide examples of active clauses with *leoten of*, which clearly show that *of* can be elliptical (cf. a verb like *leave off*) when *leoten of* means *think of* (1993: 143). Although not all aspects of Denison’s objections are sufficiently supported, his remarks question the status of the example, and especially the status of *of*, to such an extent that this cannot be taken to be the first unambiguous example.

The second thirteenth-century example is given by Denison (1993), who adds (30) to the examples from Visser (1973).

- (30) *þer wes sorhe te seon hire leoflich lich faren so reowliche wið*
 there was sorrow to see her dear body dealt so cruelly with
 c1225 (*St. Juliana* (Royal) 22.195; Denison 1993:125, his 10)

This example is from a text which, like the previous example, survives in two different manuscripts: the Royal and the Bodley, both from the early thirteenth century. Example (30) is the example from the Royal manuscript, while the PPCME2 contains the Bodley version, which has the following version of the same sentence (also already given by Denison 1981: 219, fn 36):

- (31) *þer wes sorhe to seon on hire freoliche flesch hu ha ferdan þer-wið.*
 there was sorrow to see on her dear body how they had.fared therewith
 ‘there was sorow to see on her dear body how they had fared with it’
 c1225(?c1200) (*St Juliana* (Bodley) 104.137)

This version has a finite clause beginning with complementizer *hu*, and there is no passive and also no stranding of the preposition. The idea put forward in the PPCME2 corpus description is that the Royal, which has example (30), is earlier and closer to the Latin original: the Bodley manuscript is “an expansive revision of the original text [the Royal], a translation of a popular Latin life” (Kroch & Taylor 2000, in the discussion on the Katherine group³⁰). The fact that the first manuscript has a prepositional passive, but the revised manuscript finds a different solution suggests that the prepositional passive was at that stage not yet (completely) acceptable (cf. Warner 1982: 138ff. for a similar argument concerning ECM passives, see section 3.4.2.1 for more details). Note that it is not the fact that it is a translation from Latin that is relevant here—after all, Latin does not have prepositional passives—but the revision. The revision possibly indicates that the prepositional passive was, at this stage, not yet accepted, which means that again, this example cannot be considered the first unambiguous example of a prepositional passive. What it does show, however, is that this period may represent an initial stage in the development towards a fully grammatical prepositional passive.

³⁰ <http://www.ling.upenn.edu/histcorpora/PPCME2-RELEASE-3/info/katherine-group.html> [accessed 31.05.2013]. Kroch & Taylor base most of their discussion on Savage & Watson (1991) and Dobson (1972, 1976).

The PPCME2 provides the third thirteenth-century example which could qualify as an early prepositional passive.

- (32) *Sum is ald ant feble ant is þe lesse dred of*
 Some are old and weak and are the less feared of
 c1230 (cmancriw, l.44.43-6)

Dred here has the label VAN, and if it is indeed a passive participle, the translation will be ‘Some are old and feeble and are to be feared less of’. Indeed, Savage & Watson (1991: 49) provide a similar translation of ‘One is old and weak and is the less to be feared for’. However, *dred* also occurs as a noun in ME, in which case the translation will be ‘Some are old and feeble and there is less fear of them’. Unfortunately, the other instances of *dred* in the PPCME2 do not immediately solve this puzzle: some are labelled as a noun (N) and others as a passive participle (VAN), even though the structural difference is not immediately clear, witness (33) and (34). The translations are from Savage & Watson (1991: 119 and 128, respectively), on which the parse of the text in the PPCME2 is also based (Kroch & Taylor 2000).

- (33) *Gastlich fondunge þt is mare [_N dred] of; Mei beon for þe peril icleoped breast wunde.*
 spiritual temptation which is more fear of; may be for the danger called chest wound
 ‘Spiritual temptation, which is more to be feared, can be called a chest-wound, on account of its danger.’
 c1230 (cmancriw, ll.146.1971)

- (34) *& hwæn ha eftær feole 3er feleð ham se stronge . wundreð hire swiðe & is of [_{VAN} dred] leoste god habbe hire al for warpen .*
 and when she after many years feels them so strong, wonders her very and is of feared lest god has her all away cast
 ‘And when after many years, she feels strong temptations, she is very shocked, and afraid that God has cast her away completely.’
 c1230 (cmancriw, ll.161.2217-8)

Interestingly, the noun is translated with a participle by Savage & Watson, while the participle in (34) is translated with a noun, *afraid*. More support for treating *dred* as a noun comes from Bosworth & Toller (1898), who include *ofdræd* a single word in their Anglo-Saxon Dictionary, with the meaning ‘terrified, afraid’. Denison (1985, as discussed by Denison 1993) includes an example similar to (32) in his discussion of quasi-ellipsis, treating *dred* as a noun and *of* as elliptical.

- (35) *ȝet is meast dred of hwen þe sweoke of helle eggeð to a þing*
 yet is most fear of when the deceiver of hell incites to a thing
 ‘Yet there is most fear when the Deceiver of Hell incites one to something’
 c1230(?a1200) (*Ancr.* 60a.19; Denison 1993: 142, his 87)

Following the ambiguous status both of *dred* and *of*, and the likelihood that *dred* should be interpreted as a noun, example (32) cannot be considered as an early example of a prepositional passive. This means that there are no unambiguous examples in the thirteenth century. While the ambiguous examples point to a development towards prepositional passives as marginally possible, the lack of unambiguous examples in the same period sheds doubt on the status of these three ambiguous examples as proof for the acceptability of early prepositional passives in this period.

Fourteenth and fifteenth century

Visser (1973) gives a total of 28 examples for the fourteenth century, adding his own examples to Van der Gaaf’s (1929) collection of 11. Denison (1993: 160) later discounts 7 of Visser’s and Van der Gaaf’s examples, for a variety of reasons: three because they do not really represent prepositional passives (like *leoten of* above, and examples with *run at/about* and *served withal*); one because it is wrongly attributed—not from Rolle’s *Psalter* (1340) but Hilton’s *On the Mixed Life* (Thrn, c1440); and two because they are “unsafe” according to Denison. These are given as (36) and (37).

- (36) *Waltere Was smyten þorgh wip a lance.*
 Walter was pierced through with a lance
 c1338 (Rob. Of Brunne [D: Mannyng], Chron. Pt. II, 2040)
- (37) *He þat was mast forgiven till, Mast aght to luue him wit skill.*
 He that was most forgiven to, most ought to love him with skill
 13.. (Curs. M. 14048)
- (Visser 1973: 2123, §1951; glosses mine)

Denison (1993) refers to Denison (1985) for example (36), stating that *through* has a “confused categorial status” in ME, because it can be prefixal, adverbial or prepositional (1993: 141). In (37), the uncertainty may be due to the fact that it occurs in a poem which uses end rhyme, in this example shown by *till* and *skill*, although Denison does not seem to use this argument for other examples.

The example that Denison presents as the earliest unambiguous example from the fourteenth century is not entirely without problems. It is one of Visser’s examples but with a later date for the manuscript (c1330(?a1300) instead of c1300). Note that this new date means that there is a considerable gap of about a hundred years between the alleged first examples from the early thirteenth century and this example.

- (38) *bis maiden ... feled also bi her þi | þat sche was yleyen bi*
 this maiden ... felt also by her thigh that she had.been lain by
 ‘This maiden felt then by her thigh that she had been lain with’
 c1330(?a1300) (*Arthu.&M.* (Auch) 849)³¹; Denison 1993: 125-6, his 11)

Considering the fact that *of Arthour & of Merlin* is metrical and, more importantly, uses end rhyme, the question is whether this prepositional passive is influenced by those factors: *þi* in the first line rhymes with the stranded preposition *bi* in the second line. This is not to say that all poetry should be disregarded or that this is not, in fact, an example of a prepositional passive. Like the previous examples, however, it is problematic as an early example because it does not unambiguously, i.e. without effect of rhyme or metre, show that prepositional passives are fully acceptable at this stage.

Denison has a further 7 examples whose manuscript and composition date both fall in the fourteenth century. The PPCME2 has another 8 examples with both composition and manuscript date before 1400. This means there are a total of 16 examples pre-1400 in the collection. The first unambiguous examples of prepositional passives are given in (39)-(42).

- (39) *Was neuer prince, I wene, þat I writen of fond, More had treie & tene³²*
 [there] was never a prince, I know, that I written of found, who had more
 suffering
 c1338 (Rob. Of Brunne, Chron. Pt. II, p235; Visser 1973)
- (40) *He waas piled and i-robbed, and fare wiþ as it were a þeef.*
 he was plundered and robbed and dealt with as.if (he) were a thief/as.if it
 were a thief (they were dealing with)
 a1387 (Trev. *Higd.* VII 7.ix.385.14; Denison 1993: 126, his 18)
- (41) *But by pope Sergius his pistel þat was i-sent to abbot Colfridus it is i-knowe
 þat Beda was i-sent after and i-prayed for to come to Rome for to assoyle
 questiouns þat were þere i-mevede.*
 ‘But by pope Sergius’s letter that was sent to Abbot Colfridus, it is known that
 Bede was sent after and asked to come to Rome in order to answer questions
 that were raised there.’
 a1387 (cmpolych,VI,223.1599)

³¹ Visser gives Kölbing 852 as the reference.

³² This is a complicated example because of the presence of relativization, passivization, an unusual word order, and the omission of, among other things, a finite form of *to be* to combine with *The prince ... written of*. In my opinion, the right paraphrase here is *I never found a prince (who was) written of and had more suffering*, which shows more clearly that it is a prepositional passive.

(42) *3if be wardeyns of þat 3eer ben **sent after***

If the Wardens for that year are sent for

1389 (*Lond.Gild.Ret in Bk.Lond.E* 52.52; Denison 1993: 126, his 19)

There are a further 9 examples in Denison and 4 in the PPCME2 whose composition date is before 1400 but whose manuscript date is after 1400. The following two examples illustrate the productivity of prepositional passives through their occurrence in non-standard contexts.

(43) *Than, if þou ... latis þam spill for defaute of kepyng- unarayede, unkepide, and noghte **tente to** as þame aughte for to be,- thou pleses Hym noghte*

then, if thou ... lets them spill for default of keeping – unarrayed, unkept and not tent to as them ought for to be – thou pleases him not

‘Then, if you ... let them waste away because of a lack of keeping – untidy, unkept, and not cared for as they ought to be cared for – you do not please him’

c1440(a1349) (cmrolltr,29.617)³³

(44) *when any...haves envy to þam þat es **spokyn mare gode of þan of þam**;*

when any...feel envy towards them that are spoken more good of than of them

‘when any...are envious of those who are praised more than them’

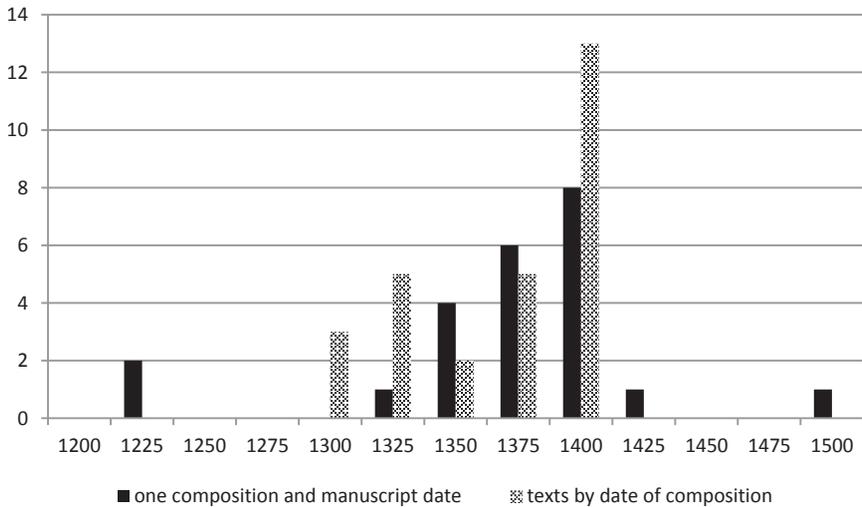
a1450(?1348) (Rolle, *Fliving* 86.43; Denison 1993: 154, his 122)

Example (43) shows a participle and preposition that seem to have been used adjectivally and example (44) is given by Denison as the first attestation of a phrasal prepositional passive, i.e. a passive of a verb which takes both a prepositional object and an adjective, such as *they speak well of them*.

Having established the first unambiguous examples, let us look at the distribution of these examples. Figures 3.2 and 3.3 provide an overview of the dates of the early examples in detail. Figure 3.2 gives the dates of composition for those examples that occur in manuscripts where there is a considerable gap between date of composition and date of the manuscript and the manuscripts which have a similar date for composition and manuscript. What this figure shows, then, is the earliest possible scenario.

³³ Visser gives the date 1340; Denison (126-127, his 20) gives a very similar example, but shortened, and provides the following reference: c1390 Hilton, *ML(Vrn)* 272.37

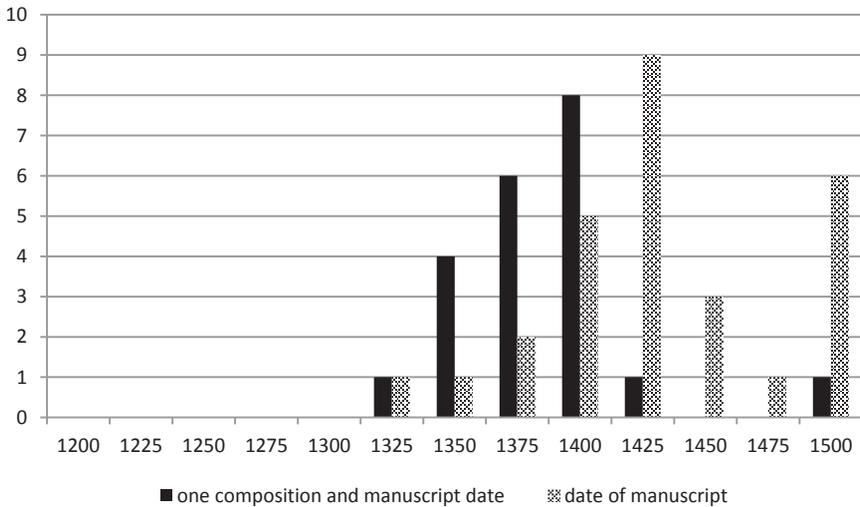
Figure 3.2 Prepositional passives before 1450, by composition date



This figure includes two of the early ambiguous examples, *yleyen bi* and *leoten of*—while excluding the third early example *dred of* because it is the clearest case of misanalysis. The figure shows that the examples in unambiguously dated manuscripts before the late fourteenth century are rare: there are the two ambiguous examples at 1225, then one example at 1338, and four between 1350 and 1375. There are a further 10 examples with composition dates before 1375, but with later manuscript dates. The figure clearly shows that the thirteenth-century ambiguous examples occur well before the prepositional passive begins to occur in reasonable numbers at all: there is a gap of more than 100 years between them and the first unambiguous example (1225-1338) and a slightly smaller gap, but still of 100 years, between the early thirteenth-century examples and the first example by date of composition (1225-1325).

Figure 3.3 presents a more careful picture of the introduction of the prepositional passive, excluding the ambiguous early examples and considering the manuscript dates instead of the date of composition. This, then, is the representation of the ‘slowest’ scenario for the introduction of the prepositional passive.

Figure 3.3 Prepositional passives before 1450, by date of manuscript



In this presentation of the introduction of the prepositional passive, the first attestations appear much later—with two examples around 1325, and a first larger clustering of examples from 1350 onwards. What remains similar to Figure 3.1 is that it is only in the late fourteenth century that they start to appear more frequently. Following these observations, I will take the middle of the fourteenth century to be the proper date of introduction of the prepositional passive—i.e. the time when it became an acceptable, grammatical construction—while allowing for the possibility that the early examples already point to a development in the direction of a prepositional passive.

A second question that is relevant to the introduction of the prepositional passive is the type of verb and preposition combinations, because these may point to a particular scenario of introduction. Table 3.1 presents the combinations per subperiod for the PPCME2 examples and Denison (1993), Van der Gaaf (1929) and Visser's (1973) collection separately. Examples which occur in both collections are given between brackets. Denison, in fact, claims he has 39 examples with composition dates prior to 1400 with 24 different verb-preposition combinations, but this does not follow from the explanation of his selection of examples.³⁴

³⁴ He discounts 7 of Visser's 29 examples (1973: §1950 and 1951) with composition dates before c1400: *leten of*, *ronnen aboute*, *seruede with-alle*, *ymediid wiþ*, *smyten þorgh*, *forgiven till*, and *lukede to*. He includes 4 examples from Visser's §1947 with P-V order: *speken of*, *senden after*, *leten of*, and *speken of*. He adds 8 examples of his own: *faren wið*, *dispensen wiþ*, *leten of* (twice), *wonder uppon*, *tenden to*, *faren wiþ*, and *leten by*. This all adds up to only 34 examples. It is possible that he has included examples from §1952 with composition dates before c1400, but he does not comment on this.

Table 3.1 Verb-preposition collocations in early examples of prepositional passives

	Composition and manuscript date before 1400	Composition before 1400, manuscript after 1400	Composition and manuscript between 1400 and 1500
PPCME2	<i>i-sent after, (snyt vpon), told by, dispensud wip, seyde of, holden vppe, in clopede, sent after</i>	<i>(Tente to), spokyn mare gode of, a-feerd of, lokid to</i>	<i>comen to (3), gouen to, spoken of, sent to (2), made sorwe fore, songen of, schryuen of, boght on, pluckyt at, ferd wyth</i>
Total PPCME2	8	4	13
Van der Gaaf/Visser/Denison	<i>faren wið, yleyen bi, writen of, trowen on, i-told of, fare wip, sent after, (tended to), (snyt upon), of spokun, after send, leten of, boght on</i>	<i>Lett of, leten bi, wondryd upon, wonderid uppon, geue to, waad ouer, told bi, dispensid wip, told by, leyn by</i>	<i>Comen to, read of, wyt don, i-set by, runnen at, lette by</i>
Total vdG/V/D	13	10	6
Total	21	14	19

With only a handful of combinations occurring more than once, there is no generalization to be made among these combinations. For instance, one combination that occurs a number of times is *sent after*, but all four instances occur in the first column. *Comen to* also has four instances in the third column, but these are all from the same text: *the Cloud of Unknowing*. If there are any generalizations to be made with respect to semantic groups, only one coherent group can be identified, the verbs of saying and thinking, making up 10 of the early examples: *seyde of*, *i-told of*, *of spokun*, *boght on*, *spokyn mare gode of*, *told bi*, *told by*, *spoken of*, *songen of*, and *schryuen of*.

The range of verb+preposition combinations shown in Table 3.1 suggests that the introduction of the prepositional passive was truly a syntactic, i.e. structural, innovation rather than one limited to one group of verbs as previously suggested by Denison (1985, 1993). He proposes a lexical diffusion account based on occurrences of *leten of/by*, *setten of/by* and *tellen of/by* with the shared meaning of ‘regard, esteem, think of’ (1993: 141). He finds that 12 of his 39 early examples (i.e. before the fifteenth century) belong to this group, and also occur at least 14 times in the fifteenth century (*ibid.* 125). His claim is that they constitute a group based both on semantic and phonological similarities, and he takes their frequency to support a scenario where the prepositional passive was first used with these verbs and later spread to other verbs (*ibid.* 141). However, the additional examples from the PPCME2 mean that this dominance of *leten* verbs is no longer clear and crucially no

longer provides enough evidence for an account of this kind in terms of lexical diffusion.

The data in this section have highlighted the ambiguous status of early examples of prepositional passives presented previously in the literature and I have argued that the earlier proposed thirteenth-century examples should not be taken as evidence that the prepositional passive was used as a grammatical construction at that stage. The first unambiguous examples date from the middle of the fourteenth century (c1338) and I will take this as the time of introduction. The fact that the prepositional passive is used with a wide range of verbs from the beginning suggests that the introduction of the prepositional passive is not a lexical innovation but truly a syntactic one.

3.2.3 A new account for the introduction of the prepositional passive

Based on the data from the previous section, this section provides a new account for the introduction of the prepositional passive, based on two principles: on the one hand, the notion of reanalysis as an important factor from the very first moment the prepositional passive was used; on the other hand, a scenario where steps of minimal alteration led to a first use of the prepositional passive and then to it becoming an accepted construction. The discussion in this section is based on all the examples from the literature and the additional examples found in the PPCME2.

3.2.3.1 Further proof for reanalysis

There are clear indications that the notion of reanalysis, the unity of verb and preposition, is relevant from the very first instances of prepositional passives onwards. The crucial aspect here is not just the theoretical notion of reanalysis, but the observation that speakers at that time treat them as units, suggesting that they perceive them to be closely connected. This section presents two pieces of evidence that they are perceived as units.

Conjunction with single verbs

The first indication that the verb and preposition are perceived as a unit is the fact that 7 of the 35 early examples, i.e. those dated around 1400, are in fact conjunctions of single verbs with verbs occurring with a stranded preposition. One example is (40) above; the other examples are given in (45)-(50).

- (45) *he* was *tormentid* and *aftyr* he was *spyt vpon*.
 he was tormented and afterwards he was spat upon
 c1400 (cmwycser,391.2966)

- (46) *and chesun hem a new weye bat mut ofte tymes be **clowtid**, and be **dispensud wip** by antecrist, as þe feend techuþ hem.*
and chose for themselves a new way that must often be clothed, and be dispensed with by Antechrist, as the Devil teaches him
c1400 (cmwycser, l, 338.2006)
- (47) *He schall be **lenede** and **louede** and **lett of** a while | *Wele more þan þe man that made it hymseluen*
he shall be listened.to and loved and thought of a while rather more [=more highly] than the man who composed it himself
c1450 (1352-c1370) (Winner and W. 27; Denison: 126, his 14)*
- (48) *Litel is he **luid** or **lete by** þat suche a lessoun techip*
'He is little loved or valued who teaches such a lesson'
c1400 (a1376) (PPI.A.(1) 11.29; Denison: 126, his 15)
- (49) *For nobley in vertues shulde be **coueytid** & worldly nobley litil **told by***
for nobility in virtues should be coveted and worldly nobility little told by [=valued]
a1500(?c1378) (Wycl. OPastor. 440.14; Denison: 126, his 16)
- (50) *Roten ordinaunce of men is **more worshipid** & **more told bi** þen þe ordinaunce of crist*
rotten government of men is more worshiped and more told by [=more highly regarded] than the government of Christ
(Wycliff, Eng. Wks. 303; Visser 1973)

All these examples have a single past participle conjoined with a passive participle+preposition combination. Example (43) above already showed another type of combination: a passive participle+preposition conjoined with adjectival participles. There are a further two examples which show a participle+preposition in close connection to a single verb but under a different configuration, given in (51) and (52).

- (51) *gif alle þe lorus ..., **Ben trewe to be trowen on** & **trusty to leue***
if all the directions ... be true to be trusted on and reliable to follow
1340-70 (Alexander & Dindimus (ed. Magoun) 829; Visser 1973)
- (52) *Vplondisshe men ... **fondeþ** ... for **to speke Frensce**, for **to be i-told of**.*
Country men ... concern themselves ... for to speak French for to be told of
'Countrymen ... attempt to speak French, in order to be held in high esteem'
c1378 (Trevisa, Higden 2, 159; Visser 1973)

In (51), *trewē to be trowen on*, an adjective and a passive infinitive in combination with a preposition, is conjoined with *trusty to leue*, another adjective but now followed by an active infinitive. In (52), there is no conjunction, but in both cases *for to*—though structurally different in the two cases, first an infinitive marker and then a purpose adjunct marker—is followed by an infinitive, first by the active *to speke Frensche*, and in the second instance by the passive *to be i-told of*. The close occurrence and their occurrence after *for to* again suggests that the single verb and the verb+preposition are treated as equal structures.

Translation of single Latin verbs

The second indication that the verb+preposition combinations are perceived to be units by the speakers are those instances where they represent translations from single Latin verbs. Examples (53)-(56) show the Latin original with the participle that is translated with a verb and preposition in English in bold as usual.

(53) =(41)

Constat tamen Bedam per epistolam Sergii Colfrido abbati **transmissam** invitatum fuisse ut Romam veniret ad enodandum aliquas nuper emersas quæstiones.

*But by pope Sergius his pistel þat was i-sent to abbot Colfridus it is i-knowe þat Beda was **i-sent after** and i-prayed for to come to Rome for to assoyle questiouns þat were þere i-mevede.*

‘But by pope Sergius’s letter that was sent to Abbot Colfridus, it is known that Bede was sent after and asked to come to Rome in order to answer questions that were raised there.’

a1387 (cmpolych,VI,223.1599)

(54) Quod quidem caelum non his potius est quam sua qua regitur ratione **mirandum**.

*The whiche hevene certes nys nat rathere for thise thynges **to be wondryd upon**.*

The which heavens certainly not.is not rather for these things to be wondered upon

c1374 (Chaucer, Boece III, pr.3, 783; Visser 1973)

(55) quod solum quanta dignum sit ammiratione **perfecto consideras**.

*The whiche thing oonly how worthy it is **to ben wonderid uppon**.*

The which thing only how worthy it is to be wondered upon

c1380 (Chaucer, Boece, IV, pr. 1, 1148; Visser 1973)

(56) torrentem quem non potui pertransire. quoniam intumuerant aquæ profundi torrentis, qui non potest **transvadari**. (Clementine Vulgate).

*The depe watris of the stronde hadden wexe great, that may not **be waad ouer**.*

‘The deep waters of the stream had waxed greatly, so that they could not be waded through.’

c1380 (Wycliff, Ezek, 47, 5; Visser 1973)

In all cases the Latin participle is translated with a passive participle+preposition in the ME text. Note that these examples, like the examples of conjunctions with single verbs, are among the early examples, all dated in the late fourteenth century. This shows, again, that the reanalysis is present from the early examples onwards and is therefore a relevant factor in the introduction of the prepositional passive.

Discussion

The number of examples that bear indications of reanalysis, either because a verb+preposition is shown in conjunction with a single verb, or because the verb+preposition unit represents the translation of a single Latin participle, is considerable, with 14 examples among the 35 occurrences in texts which have a composition date in the 1400 century (columns 1 and 2 in table 3.1 above). It suggests that from the earliest use onwards the notion of a close connection between the verb and preposition is visible, and a determining factor for the use of the passive. On the basis of these findings, we can conclude that reanalysis does indeed take place in those passives, which means that the change we need to account for is this new possibility. The conclusion that the locus of the innovation is the new structure that is the result of a reanalysis in prepositional passives means that an earlier proposed scenario from Lightfoot (1981) and Van Kemenade (1987) in terms of the loss of inherent case is no longer sufficient to account for the introduction. They propose that prepositional objects cannot be passivized in OE because that would mean movement of an inherent case-marked object (marked by the preposition) to a structural case position, the subject position, which results in a “case-clash” (1987: 211). When inherent case is lost, this problem no longer exists and any object is free to move to the subject position. However, under a reanalysis account, the NP is assigned case (or the case feature is valued, in more recent terminology, cf. Platzack 2005 and Pesetsky & Torrego 2007) by the verb+preposition unit, i.e. it is assigned structural case and can move to a structural case position without problems. This makes case in the introduction of the prepositional passive only a secondary or promoting factor, but not the crucial change. The question that remains, then, is why speakers would begin to reanalyse the verb and preposition so as to allow passivization of what was the complement of the preposition.

3.2.3.2 A new proposal: reanalysis and minimal alterations as keys to the introduction

In this section I outline a scenario which sees the introduction of the prepositional passive as a gradual process of increasing acceptability, following minimal

alterations. This scenario builds on the OE system of stranding in relative clauses and the ambiguity of perfect and passive constructions in OE and early ME.

Minimal alterations

First of all, consider the following examples. These are among the earliest examples of prepositional passives and are not translations from Latin and not conjoined with single verbs. In other words, there is no direct indication that reanalysis plays a role here.

(57) =(39)

*Was neuer prince, I wene, bat I **writen of fond**, More had treie & tene*
 ‘There was never a prince, I know, that I found written of.,...’
 c1338 (Rob. Of Brunne, Chron. Pt. II, p235; Visser 1973)

(58) *And be commaundment ys brokun, | And be halyday, byfore **of spokun**.*

‘The commandment is broken, and the holy day previously spoken of.’
 a1400 (c1303) (Mannyng, HS 1033; Denison 1993: 126, his 12)

In these two examples, the prepositional passive occurs in a relative clause. This is not merely a coincidence, I believe, but an extremely relevant factor. Stranding of prepositions in relative clauses was already possible in OE, but crucially not in passives (see Allen 1980 and Van Kemenade 1987). Regular relative stranding is illustrated in (59).

(59) *On þam munte Synay, þe se Ælmihtiga **on becom**, wearð micel ðunor gehyred*

on the mountain Sinai which the Almighty on came, was great thunder heard
 ‘On mount Sinai, on which the Almighty came, great thunder was heard’
 (ÆCHom II, 12.1.116.226; Fischer et al. 2000: 66)

The preposition *on* is stranded in the relative clause (but note the P-V order), similar to examples (57)-(58) above. In those early passive examples, then, the stranding may be allowed (licensed) by the fact that the NPs sentences are relativized, a process which is unrelated to their being passivized. What we also find in the same period are examples of passives occurring in relative clauses, which look very similar to prepositional passives, but where the stranding is due to relativization, not passivization.

(60) *for noman to him hade suspesion, for enchesoun of his habit bat he was **in clopede**,*

for nobody to him had suspicion, for reason of his habit that he was in clothed
 ‘For nobody suspected him, because of the clothes that he wore,’
 c1400 (cmbrut3,63.1884)

In this example, there is a stranded preposition, but the subject of the relative clause, *he*, is not the complement of *in*; instead, the complement of *in*, *his habit*, has been relativized. The occurrence of the passive in either a relative clause, or a combined clause with relativization and stranding due to passivization, then, may be an example of a minimal alteration: the passive with stranding is possible only because the stranding is less noticeable due to the relative clause.

Another observation is that many of the early examples above have the order of preposition before verb. This makes them similar to two types of OE compound verbs which consist of a verbal root and a prefix or particle: Inseparable Complex Verbs as in (61) and Separable Complex Verbs as in (62).

- (61) *Witodlice ic dyde þæt þa gewurdon cristene ealle in ðam cwarterne be-*
clysoðe wæron, ...
 truly I did that they became Christians all who in the prison in-shut were
 ‘Truly, I brought about that those shut up in prison all became Christians,...’
 (coaelhom, ÆHom 24:170.3871; Los et al. 2012: 144, their 1b)

- (62) *And seo helle þone deofol ut a-draf*
 and the hell the devil out PREF-drove
 ‘And Hell drove out the devil’
 (conicodC, Nic [C]: 282.274; Los et al. 2012: 145, their 8)

The Separable Complex Verbs, which later develop into the PDE Verb Particle Combinations, generally occur with P-V order in OE, while they become “dominantly postverbal” in ME (Los et al. 2012: 140-3). Another interesting similarity is that the particle can be stranded by movement of the verb, both in OE and ME, as illustrated in (63).

- (63) *Strupeð hire steort-naket, and heoueð hire on heh up*
 strip her stark-naked, and lift her on high up
 ‘Strip her naked, and lift her up high’
 (cmmarga, 84.471; Los et al. 2012: 143, their 7b)

The similarity to the Separable Complex Verbs in order of P-element and verb, as well as the close structural relation between particles and prepositions—they are historically related to adverbs and prepositions (Booij and Van Kemenade 2003, Los 2004)—is another context which could be proof of a minimal alteration, due to the ambiguity between the different types of P-elements, further supported by the fact that stranding of particles is also possible. Moreover, the similarity to these Compound Verbs (both Separable and Inseparable) adds further support for the notion that verb and preposition can form a unit, in the sense that this idea is not a foreign one to speakers of English in that period. The order of P-V, similar to that in SCV which would allow stranding in relatives, makes the stranding in a *passive* less obvious.

The last issue that contributes to an idea of minimal alteration is the ambiguity between perfect and passive in OE. As Los (1999) explains, there are two possible interpretations for causative/ergative verb pairs such as the one illustrated in (64), either passive or perfect.

- (64) *þa wæs an gereord on eallum mancynne. And þæt weorc wæs begunnen ongean godes willan*
 ‘then was one language among all mankind and the work was begun against god’s will’
 ‘At that time all people spoke one language, and the work had been begun against God’s will’
 (ÆCHom I, 22 318.17)

(Los 1999: 103, her 11)

This example can have both an ergative perfect interpretation (*the work had begun*) and passive (*the work had been begun*). Now consider the examples in (65)-(67), from O3 (950-1050).

- (65) *ic sceal ærest afyllan þa þincg þe ic fore asend eom*
 I must first fulfil the things which I for sent am
 ‘I must first fulfil the things for which I am sent’
 (coalive,ÆLS_[Abdon_and_Sennes]:118.4791)

- (66) *in to þære arke þe he ær of gesend was.*
 in to the Arc that he earlier of sent was
 ‘into the Arc that he was earlier sent of’
 (coadrian,Ad:21.2.51)

- (67) *þone ilcan welegan mon, þe heo ær from sended wæs,*
 the same wealthy man, who she earlier from sent was
 ‘the same wealthy man, who she was earlier sent from’
 (coblick,LS_25_[MichaelMor [BlHom_17]]:199.47.2548)

These are all perfects, and relatives with stranded prepositions. As we have already seen, stranding in relatives, regardless of whether the verb is a passive or not, is no innovation, and is already possible under the existing stranding possibilities. What is relevant here, however, is the ambiguity between passive and perfect, allowing for a passive reading with stranding. As in the previous examples, the fact that stranding is not generally acceptable in passives could be obscured by the relative clause and the perfect reading. Again, this provides an indication that possibilities for stranding became blurred because of minimal alterations, in the previous case because of stranding in relative clauses, in this case because of stranding in relative clauses and on top of that the ambiguity between perfect and passive. With the distinctions blurred, and structures occurring in the language that are close to prepositional

passives, this provides a situation in which the possibility arises for full prepositional passives to occur without much notice, in turn creating the possibility for increasing acceptability for prepositional passives among speakers of the language.

A note on Dutch

There is suggestive evidence from Dutch, a V2 language with both Separable and Inseparable Compound Verbs (Los et al. 2012) and more restrictions on preposition stranding, especially in passives, that prepositional passives are in some cases marginally allowed. Schippers (2012) includes prepositional passives in a study of acceptability judgments of preposition stranding in Dutch, through a questionnaire filled out by 65 adult native speakers of Dutch (2012: 125). Examples (68)-(70) illustrate the test sentences for preposition stranding in passives.

- (68) *De meeste ouderen worden weinig mee gepraat.*
the most elderly become little with talked
'Most elderly are hardly being talked to.'
(Schipper 2012: 219, her T14)

- (69) *#[Alleen serieuze biedingen]_i worden [PP op t_i] in gegaan.*
only serious bids become up in gone
'Only serious bids will be responded to.'

- (70) *#[Alleen bovengenoemde regels]_i werden [PP naar t_i] gekeken.*
only above-mentioned rules became to looked
'Only the above-mentioned rules were looked at.'
(*ibid.* 113, her 12a-b)

Schippers reports acceptability scores of between 2.7 and 3.1 on a 7-point-scale, with example (68) as the highest scorer (*ibid.* 130). Schippers concludes from this that pseudo-passives are "(marginally) acceptable" (130). With this score, they are slightly less acceptable than stranding in left-dislocations, as in example (71).

- (71) *Aardbeien houd ik van.*
strawberries love I of
'Strawberries I love.'
(*ibid.* 124, her xii)

Opinion seems to be more divided about these items, because while they receive a score as high as 4.1, indicating a higher acceptability, the lowest score is as low as 2.4.

The observation that prepositional passives marginally occur in Dutch, or alternatively are not considered completely ungrammatical, is relevant for the scenario for the introduction of the prepositional passive in English, because this could be evidence of a first step in a minimal alteration scenario. What it shows is

that a construction like the prepositional passive can develop in a language in the margins. If it is then promoted by other factors, it is likely to become more acceptable and thus more frequent. There are crucial differences between Dutch and English that could explain an increasing acceptability in English—notably in this regard the increasing possibilities for preposition stranding in English, the context of relative clauses for many early passive examples, and the ambiguity in status and position of the P-elements. Chapter 5 will show another of the promoting factors that English has but Dutch lacks: the increasing demand for subject strategies.

3.2.4 Summary and conclusions

This section on the prepositional passive aimed to answer the questions of the time of introduction of this type of passive and to provide an explanation for this introduction in the ME period. I have argued that the introduction of the prepositional passive must be dated in the early fourteenth century rather than the thirteenth century, as proposed earlier. I also argued that the introduction of the prepositional passive was due to very specific contexts available in English at that time. The possibilities for stranding allowed for a scenario based on minimal alterations and it is telling that the first prepositional passives that are not translations or conjunctions occur in those contexts.

The study of the prepositional passive in the PPCME2 provided additional examples to those given in the literature before this annotated corpus had become available, shedding new light on the frequency and properties of the prepositional passives in the fourteenth century. Previously, authors such as Van der Gaaf (1929, 1930), Visser (1973) and Denison (1985, 1993) had presented thirteenth-century examples of prepositional passive, but I have shown that these are ambiguous for reasons such as influence of rhyme, incorrect translations, or confusion about the manuscripts they occur in. I therefore concluded that these cannot be taken as genuine first examples of prepositional passive, with the added observation that the gap between these early ambiguous examples and the first unambiguous example is so large that it is unlikely that those ambiguous examples provide evidence that the prepositional passive was already acceptable. Whether early examples are taken into account or not, the observation that remains is that the prepositional passive only starts to become more frequent in the late fourteenth century.

One of the themes that proved to be relevant for the prepositional passive was the notion of reanalysis: the idea that verb and preposition form one unit. While I proposed an analysis of the prepositional passive in terms of this reanalysis, previously shown to be relevant by authors such as van Riemsdijk (1978), Hornstein & Weinberg (1981), the crucial aspect of my analysis was the role which reanalysis played from the first instances onwards and I have used the reanalysis as an explanatory factor, connecting PDE observations to the OE data. This, in combination with contexts where the prepositional passive was obscured because it occurred in relative clauses, where stranding was already possible, formed the basis for a different proposal to explain the introduction of the prepositional passive: it first occurs in contexts of minimal alterations where the combination of stranding

and passive is blurred by other factors, such as relatives that already allow stranding, and afterwards it slowly gains ground in other contexts as well beyond the relative clauses. Crucially, the change that has taken place is explained with reference to the context of English syntax in the period under discussion which facilitates the reanalysis. The question whether the prepositional passive continued to increase will be addressed in Section 6.3.1, which presents a study of the occurrence of the prepositional passive after the ME.

Section 3.3 The introduction of the recipient passive

This section investigates the introduction of the recipient passive in the fourteenth and fifteenth century, and aims to determine the time of and reason for the introduction of this new passive. It discusses the earliest examples of recipient passives, based on previous accounts in the literature and complemented by a corpus study of the PPCME2. In addition to the recipient passive, the corpus study also includes related passives that play a role in the introduction of the recipient passive, most notably the theme passive and the dative-fronted passive. Based on the findings from the corpus study, I argue that it was the loss of the dative-fronted passive—in all likelihood part of the loss of V2—that led to the introduction of the recipient passive.

In early syntactic work on the recipient passive, authors proposed that the recipient passive, as in *She was given the book*, was introduced in the early thirteenth century (Van der Gaaf 1929 and Visser 1973), but Allen (1995) later argued that the introduction proper took place in the late fourteenth century. At the time that the recipient passive was introduced, two other passives formed from ditransitives with two NP objects were used in English: first, the dative-fronted passive (*Her was given the book*), which was used in English up to the early fourteenth century; and second, the theme passive (*The book was given her*), where it is the direct object (theme) rather than the indirect object (recipient) that is passivized from a ditransitive active with two NP objects. This second type of passive was lost in standard English (Allen 1995), although it is still used in some British English dialects (see Siewierska & Hollman 2007 for a discussion). The recipient passive is a specifically English phenomenon, since no other West-Germanic languages allows its formation; Dutch and German, for instance, have the dative-fronted passive but no recipient passive (Baker 1988).

Most accounts of the introduction of the recipient passive build on a reanalysis having taken place, but reanalysis of a crucially different kind from the type that played a role in the development of the prepositional passive. In the case of the recipient passive, it is not a reanalysis of two elements as one unit, but rather of the grammatical status of arguments in a clause: on the one hand, there are proposals that the preverbal position is reanalysed as a subject position in the fourteenth and fifteenth century and so any element occurring before the verb must occur in nominative case (e.g. Van der Gaaf 1929); on the other hand, it has been suggested that the indirect object is reanalysed as a direct object, and consequently this object

could be passivized like other direct objects (e.g. Lightfoot 1981, Mitchell 1985). Interestingly, most of these proposals, from both perspectives, state similar factors as causes for this change, namely the loss of case marking and the increasingly fixed SVO word order in English.

This section is structured as follows. Section 3.3.1 provides a definition of the recipient passive in PDE, both in terms of the types of related constructions that are relevant to its use both in PDE and in earlier stages, as well as a syntactic analysis. The analysis builds on the notion presented in Chapter 2 that the passive targets the immediately postverbal NP, whatever its semantic role, in turn providing the basis for an account of the changes that took place in ME. Section 3.3.2 discusses the introduction of the recipient passive on the basis of previous literature (because no earlier examples were found in the PPCME2), both the earliest examples as well as earlier accounts for the introduction. Section 3.3.3 proposes that it is the loss of the dative-fronted passive—supported by other changes such as the loss of case marking and fixation of SVO order that took place at the same time—that is the cause of the introduction of the recipient passive, based on the corpus study of the recipient passive in the PPCME2, and building on the analysis of the recipient passive presented in 3.3.1.2. Section 3.3.4 concludes.

3.3.1 Defining the recipient passive

Several constructions are available in English that are related to the recipient passive *She was given the book*, both active ditransitives as well as different types of passives that are formed from these. One such construction is the theme passive, which can be formed from the same active (*He gave her the book* > *The book was given her*); another relevant type construction is the passive with prepositional phrase recipients (*He gave the book to her*; *The book was given to her*). The theme passive had been present in ME before the introduction of the recipient passive and the PP passive was at that time beginning to become established (Allen 1995). Syntactically, two questions are relevant: (i) is it really an indirect object that is passivized, as seems to be the case on the surface? and (ii) is there a single analysis that can account for all different possibilities of passivization of ditransitive actives, given that most analyses do not opt for a unified account here (e.g. Burzio 1986, Platzack 2005)?

3.3.1.1 The recipient passive and related constructions in Present-day English

Two types of active ditransitive clause types are relevant to the development of the recipient passive. The first type of active ditransitives are those that take two NP objects, one the RECIPIENT and one the THEME, as in (72). In this and the following

examples, the verbs are given in bold, the RECIPIENTS are in double underlined, the THEMES are single underlined.³⁵

- (72) [Angelica Garnett's] creativity must have been partly genetic, for she was the daughter of the artists Vanessa Bell and Duncan Grant, a parentage that **gave** her a double share of Bloomsbury inheritance.
(www.guardian.co.uk, 7 May 2012)

In this example *her* is the RECIPIENT (or GIVEE), *a double share of Bloomsbury inheritance* is the THEME (or GIFT), and *a parentage* is the GIVER. The most common type of passive that is formed from these sentences is the recipient passive, as illustrated in (73)a and (73)b

- (73) a. Labour MP Tom Watson has suggested that News Corporation **was given** details of private discussions about the culture select committee's controversial phone hacking report, in an escalating row with his Conservative counterpart Louise Mensch.
(www.guardian.co.uk, 3 May 2012).

b. In 1837, the elder Hong had, over the course of 40 days, been gripped by visions in which he ascended to a "beautiful and luminous place" where he **was given** a sword to kill demons.
(www.newyorktimes.com, 30 March 2012)

A second passive can be formed from a ditransitive with two NP arguments, witness (74)a and (74)b, which show passivization of the THEME. Its use is much more restricted than the recipient passive—see Huddleston & Pullum (2002: 249), who state that the recipient passives are "quite strongly favoured"; see also Siewierska & Hollmann (2007) and references cited there.

- (74) a. The cushion on my chair **was given** me by my daughter Emily, as was the spoon and fork mobile that hangs behind my head, startling me with its sweet chime when I collide with it.
(www.guardian.co.uk, interview, 13 April 2007)

b. When I went to Oxford I was surprised to find young men of my own age, 18, having pronounced views about food. They didn't like this, they didn't like that, whereas I just ate what **was given** me.
(www.nytimes.com, interview, 16 September 1984)

³⁵ Note that the recipient passive is also referred to as the indirect passive in the literature (e.g. Visser 1973, Denison 1993), or as the 'first passive' (Huddleston & Pullum 2002), but I will follow Allen (1995) in calling it a recipient passive.

It is important to note that the recipients in these two sentences are pronouns; as Siewierska & Hollmann (2007) point out in their study of theme-recipient order and its passives in the Lancashire dialect, pronominal arguments allow for a greater variety of word orders. It should also be noted that next to the rare theme passive, there are also dialects which have an additional rare active, as in (75).

- (75) a. He **gave** the book Mary/her.
 b. He **gave** it her.

As both Siewierska & Hollmann (2007) and Huddleston & Pullum (2002: 248) point out, this is a dialectal variant (specifically in Britain, according to Huddleston & Pullum). Indeed, some authors claim that this type of active is ungrammatical, either in British or American English (e.g. Baker 1988, Larson 1988).

As extensively discussed in the literature, the active in (72) alternates with another type of active, which has the same arguments and the same meaning, illustrated in (76). The order of RECIPIENT and THEME is reversed, and the RECIPIENT is realized as a PP, rather than an indirect object NP (see Bresnan et al. 2005 and Siewierska & Hollmann 2007 for discussion and references).

- (76) An 83-year-old man has become the oldest living kidney donor in the UK and the oldest person in the country **to give** a kidney to a stranger.
 (www.guardian.co.uk, 17 May 2012)

There is only one type of corresponding passive that is possible for this active, illustrated in (77)a and (77)b below.

- (77) a. Zopiclone **was given** to more than 5.2m patients nationally, making it the most popular sleeping tablet.
 (www.guardian.co.uk, 11 May 2012).
 b. The Kerrs expect about 40 partners to be at Aqueduct on Sunday, as well as a good-size group of people who have helped along the way, like veterinarians, farriers and the trainers who prepared Notinwildestdremz for racing before she **was given** to Turner in January.
 (www.nytimes.com, 19 April 2012)

While there also exist other types of verbs which take two internal arguments, such as *force* in (78), which takes an NP object *Google* and a clausal constituent *to back down*, these will not be discussed here but be left to see section 3.4 because of their relevance to ECM passives.

- (78) Never shy of political calculations, Hollande seized the occasion to be seen as the one who **forced** Google to back down.
 (www.guardian.co.uk, 4 February 2013)

Table 3.2 presents an overview of the actives and passives in PDE discussed in this section. To sum up, two alternations are relevant—one is the expression of the recipient either as an NP or a PP; the other is whether the recipient and theme are pronominal or nominal elements, with pronominal objects allowing for more word order variation than nominal objects.

Table 3.2 Overview of active and passives ditransitives in PDE

Actives	Passives
John gave the book to Mary	The book was given to Mary
John gave Mary the book	Mary was given the book
%John gave the book Mary	%The book was given Mary

3.3.1.2 A syntactic analysis of the recipient passive

Having outlined the grammatical possibilities in PDE, this section turns to the syntactic analyses of these different types of ditransitive actives and their corresponding passives. A traditional problem faced by the analysis of the recipient passive is that despite the common generalization that the passive targets direct objects only (e.g. Burzio 1986, Siewierska 1984), in this case it seems that the indirect object is passivized. Another problematic factor is that theme and recipient passive need to be accounted for in one analysis, in particular for those dialects or languages that allow both options. As proposed in Chapter 2, I take the passive to target the immediately postverbal NP rather than the direct object (i.e. not an analysis in terms of Case or types of objects), which means the problem of it being the ‘indirect object’ that is passivized no longer exists. This still leaves unresolved, however, why sometimes the recipient and sometimes the theme is passivized.

Early analyses of the double object construction, such as Chomsky (1981), Burzio (1986) and Larson (1988), build on differences in Case assignment between direct and indirect objects, and consequently account for the (un)grammaticality of the theme and recipient passive with reference to Case as well. Structural Case is assigned to direct objects by the verb or a preposition and inherent Case is traditionally only assigned to indirect objects (see Chomsky 1981: 170-1 for an explanation, but also note that Chomsky takes a different perspective, with an additional *V'* layer). In Larson’s proposal, the grammaticality of the recipient passive and ungrammaticality of the theme passive is due to his definition of the passive operation which applies only to inherent Case, and to his proposal that the active *Someone sent Mary a letter* is already a derived structure, from *Someone sent a letter to Mary*. Burzio (1986: 186-187) proposes that the order of elements in the derivation determines which element is passivized, i.e. in *Someone sent Mary a letter*, the indirect object *Mary* is passivized. The theme passive is ungrammatical because the indirect object fails to receive Case. Burzio’s and Larson’s accounts both build on the assumption that *A book was given Mary* is ungrammatical (or in

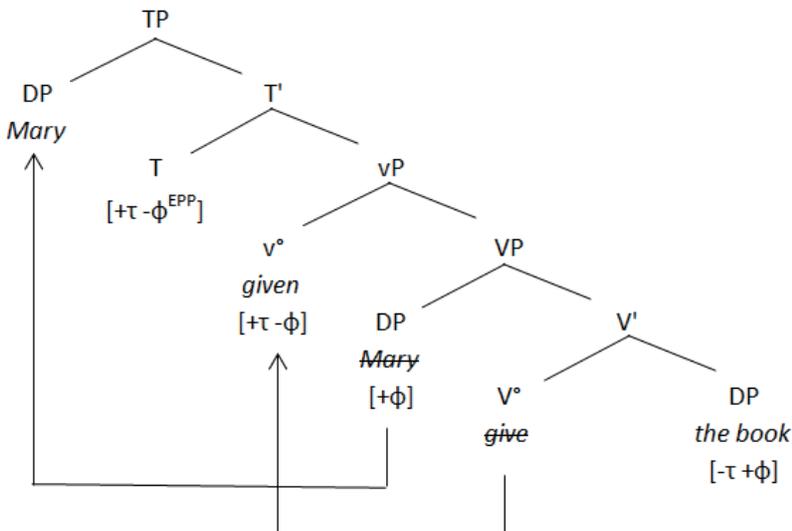
Larson's case "marginal" (363)), which is not helpful for an attempt to account for the changes in the history of English, with ME allowing for both options (Allen 1995). More importantly, it is not immediately clear how a change could come about in these systems, unless the basics of case assignment are changed or the nature of objects changes, as indeed Lightfoot has argued in (1979) and (1981) respectively.

A more recent account is presented by Platzack (2005), who focuses not only on English but also aims to account for the different possibilities for passivization of ditransitive verbs across Germanic languages. He compares English, which only allows the formation of a recipient passive of the active in (79) as in (80), to languages like German and OE, which only allow the theme passive in (81).

- | | |
|-------------------------------|-------------------------|
| (79) John gave Mary the book. | He gave her the book. |
| (80) Mary was given the book. | She was given the book. |
| (81) The book was given Mary. | The book was given her. |

Platzack proposes that the recipient and theme passive are based on the same active structure in (79). He presents an account in terms of the Probe-Goal-Agree system (see Section 1.3), and provides considerable detail as to the features involved and their status as valued or unvalued features. His analysis for the recipient passive as it occurs in PDE is illustrated in Figure 3.4.

Figure 3.4 Platzack's (2005: 6, his 10) analysis of the recipient passive in PDE



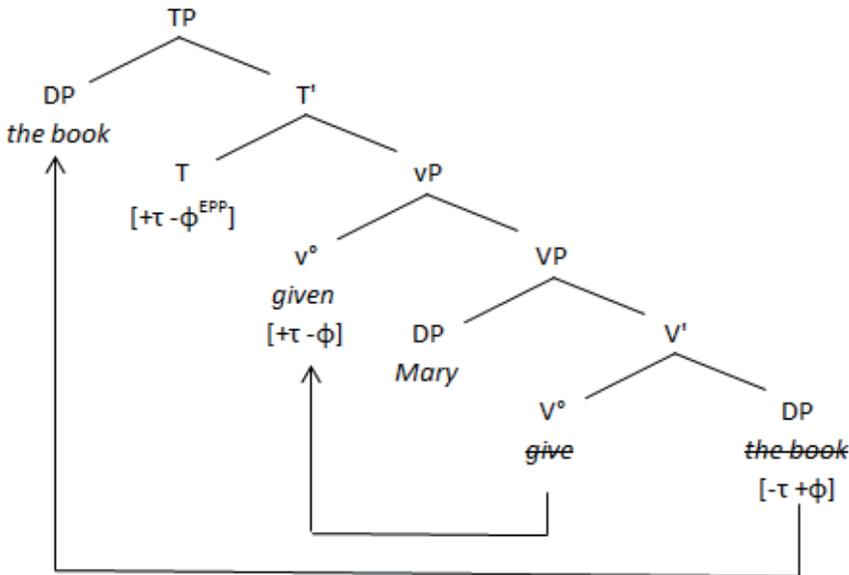
Platzack proposes that the indirect object lacks the uninterpretable³⁶ feature [τ], which he takes to mean, essentially, that the indirect object lacks structural Case

³⁶ Uninterpretable means that it is a feature which must be checked and deleted in the derivation.

and is in that way different from the direct object. The derivation of the recipient passive takes place in two steps. First, v° , which has a valued τ -feature and an unvalued ϕ -feature, looks down the tree for a match and finds *the book* as the only element with matching features because *Mary* does not have a τ -feature. The unvalued ϕ -feature on v° is valued by *the book*, while the unvalued τ -feature on *the book* is valued by v° , leading to a deletion of these features. Second, the head T, which has an unvalued ϕ -feature and an EPP-feature (demanding movement of an element to the Spec,TP position), looks down the tree for a matching element and finds only *Mary* with a matching, unvalued ϕ -feature. The feature on T is valued and *Mary* moves to Spec,TP.

Platzack's assumption that the two types of passives are based on a similar type of active means that the theme passive, as used in Dutch and some British dialects, presents something of a puzzle, because it is the *the book* (the lower NP) that moves to Spec,TP and not *Mary*. The option that Platzack (2005: 7) chooses to account for this is to assume that *Mary's* features are invisible to T. The only available argument, then, is *the book*, which has matching features. Again because of T's EPP feature, *the book* moves to Spec,TP. Figure 3.4 illustrates this derivation.

Figure 3.4 Platzack's (2005: 7, his 12) analysis of the theme passive



The first step in the derivation of the theme passive is similar to that in the derivation of the recipient passive: the features on v° and *the book* are valued against each other. In the second step, T looks down the tree and because *Mary* does not have any features, it now only finds *the book*, which still has matching

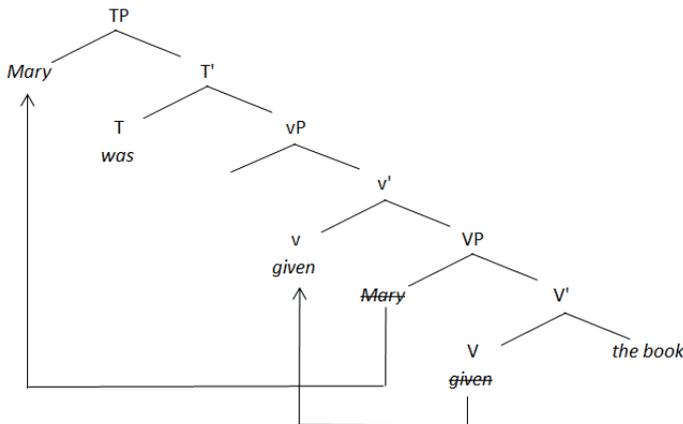
features.³⁷ Because of the EPP-feature on T, the matching element, *the book*, moves to Spec,TP. These two analyses, Platzack argues, account for the differences between Modern English and OE and German (which still have dative case). Platzack's Minimalist account indeed allows for both passives, but, like the previous accounts, it builds on a stipulated change in the nature of the indirect object.

I would like to follow up the analysis of the passive targeting the immediately postverbal NP, as presented in Section 2.3.3, by proposing that the theme and the recipient passives are, unlike in Platzack's analysis, the passives of two distinct active constructions.³⁸

- | | |
|----------------------------|------------------------------|
| (82) <i>Active</i> | <i>Passive</i> |
| a. John gave Mary the book | Mary.NOM was given the book. |
| b. John gave the book Mary | The book.NOM was given Mary |

In both examples, it is the immediately postverbal element that is passivized, as the closest available argument. This means that *Mary* in the first example is not grammatically an indirect object, even though it is understood as having a 'recipient' or 'goal' θ -role. This is along the lines of Burzio (1986: 186-187): it is the structure that is relevant, not the θ -roles or the notion of indirect object. The following figures illustrate the resulting derivation for the theme and recipient passive.

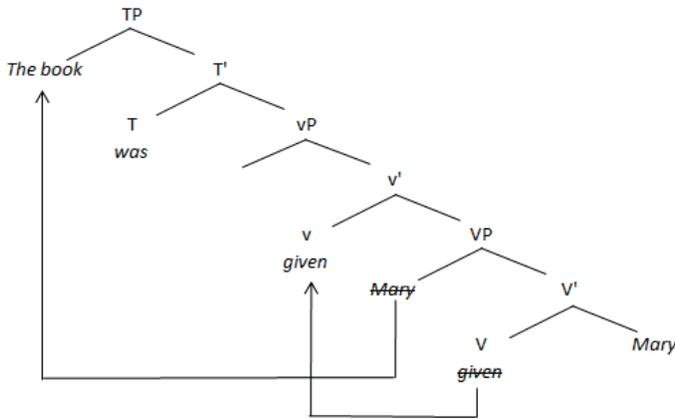
Figure 3.5 The recipient passive



³⁷ The features on *the book* have already been valued by v° . It is not clear to me whether (or which type of) valued features can twice in one derivation provide a value for unvalued features. If it is indeed the case that the ϕ -feature on *the book* can provide a value both to v° and T, then why does T in the recipient passive not choose *the book* as its goal, which is a better match in terms of full features than *Mary*?

³⁸ Larson (1988: 362) also states that the passives are based on two different actives and that the passive targets the immediately postverbal object. However, his active is already a derived structure (*John gave the book to Mary* > *John gave Mary the book*), which is different from what I propose.

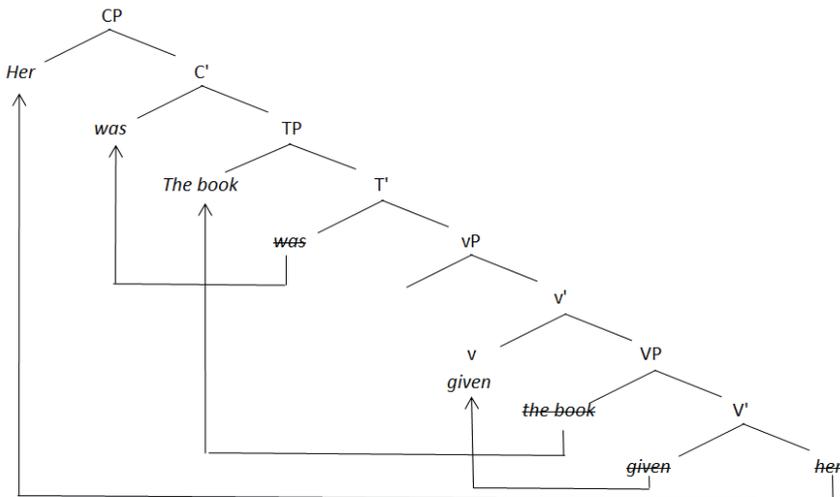
Figure 3.6 The theme passive



The highest internal argument, in Spec,VP—*Mary* in Figure 3.5 and *the book* in Figure 3.6—is the closest available argument from the perspective of T, because no external argument is merged in Spec,VP. The argument in Spec,VP moves to Spec,TP to satisfy T’s EPP feature. When the recipient is expressed as a PP (*John gave the book to Mary*), the analysis is straightforward: the NP object is the only available argument to move to the subject position.

A consequence of this analysis for the level of language use is that it suggests that passives are only possible if the corresponding active is possible, although I would like to argue that this parallelism only holds to a certain extent, because it is not true generally that all passives have an active equivalent in the language (cf. Section 2.3.1). For Standard English, this prediction seems to be borne out: it has the active and passive pair in (82)a but neither version in (82)b. As previously discussed, some dialects allow *The book was given Mary*, and some dialects allow *He gave the book Mary*; but I do not know whether this correlation is complete for each dialect (and perhaps this is not necessary if some versions are archaic forms). The other West-Germanic languages seem to present a contradiction with respect to this prediction: Dutch and German have the active corresponding to *John gave Mary the book*, but only allow passivization of *the book*, which would be the version corresponding to *John gave the book Mary*. However, what these languages do have as a seemingly corresponding passive to *She was given the book*, is the dative-fronted passive, like OE. I will assume that the dative-fronted passive is a theme passive with a fronted recipient, as illustrated in figure 3.7 below (abstracting away from some details).

Figure 3.7 The dative-fronted passive in German and Dutch



In the dative-fronted passive, based on *John gave the book Mary*, first the direct object *the book* is passivized, and then the indirect object *Mary* is fronted to Spec,CP.

The analyses presented here for the recipient passive, theme passive and dative-fronted passive, which build closely on the available active structures, mean that two questions need to be answered for the introduction of the recipient passive: (i) what are the active structures available in ME and do they form the basis for the available passives? (ii) under this analysis, why is the recipient passive not possible before ME? Before I continue to outline a scenario for the introduction of the recipient passive based on these syntactic analyses and the data from the PPCME2, the next section presents an account of the time and reason for the introduction of the recipient passive according to previous researchers.

3.3.2 Previous accounts of the introduction of the recipient passive

This section discusses the first examples of the recipient passive that have been presented in the literature, followed by a discussion of previous accounts given for the introduction of the recipient passive. These accounts are all based on reanalysis, either of the preverbal argument as the nominative subject, or of the postverbal argument(s) as a direct rather than indirect object following the loss of distinction between direct and indirect objects after the loss of case marking on nominal arguments.

3.3.2.1 Earliest examples

Allen (1995: 389-95) reviews in detail the earliest examples of the recipient passive given in the early works on early English syntax—Jespersen (1927), Van der Gaaf

(1929), and Visser (1973). All three authors present examples dating from before the middle of the fourteenth century, but Allen discards most of these early examples. One of the main reasons to reject examples is because of what she sees as various forms of misanalysis, starting with (83) and (84).

- (83) *þa he was þus ilete blod*
when he was thus let blood
'When he was thus let blood'
(AW 61.24; Allen 1995: 389, her 247)

- (84) *and some were cut the hals*
and some were cut the neck
'And some had their necks cut'
(Ch. LGW. 292; Allen 1995: 390, her 248)

Allen (1995: 390) excludes these examples—following Jespersen (1927: 300, 312), although the examples are still given in Van der Gaaf en Visser—because she considers *blood* and *hals* not to be true direct objects. An important piece of evidence is that these objects cannot passivize: *he was let blood* is attested, but *blood was let him* is not. She also notes that these types of verb and object combinations already allowed what she calls 'personal passives', i.e. with the person as subject, long before recipient passives were grammatical.

Another type of misanalysis, according to Allen, is represented by example (85), dated around 1300, from Van der Gaaf (1929).

- (85) *Swyche a man þat þus ys shryue, May be assoyled and penance 3yue*
such a man that thus is shriven may be absolved and penance given
'Such a man that is shriven thus may be absolved and given penance'
(H. Synne 10854; Allen 1995: 390, her 250)

Allen gives an alternative interpretation to the translation given in (85): 'the man can be absolved, and penance can be given', with *penance* as the subject, instead of *Jesus* (1995: 391), which means that the sentence no longer has a recipient subject. Another example from Van der Gaaf, given in (86), should also be excluded because it has been given the wrong date (it is not early fourteenth but late fourteenth century).

- (86) *I fand hesu bowndene, scourgede, gyffene galle to drynke...*
I found Jesus bound, scourged, given gall to drink...
'I found Jesus bound, scourged, given gall to drink...'
(Hampole Pr. Treatises p. 5; Allen 1995: 390, her 251)

At the same time, Allen gives (87) as an early and similar example to (86), showing that the same type of passive is in fact already found in the early fourteenth century.

- (87) *Oðer gef me is iluuet mare þen an oþer mare iolhnet mare idon god oðer menske*
 Or if one is loved more than an other more caressed more done good or honor
 ‘or if one is loved more than another, more caressed, benefited or honored’
 (AW 93.26; Allen 1995: 391, her 253)

Allen states that it cannot be proven that (87) represents an innovation, saying that “one would have to demonstrate that such examples are not found in OE writings, and I do not know whether this is so or not” (*ibid.* 392). She does state, however, that *do* did not occur with recipient passives at this time yet and she ultimately concludes that it is parallelism with the preceding two passives in the sentence, *loved more* and *more caressed*, that has led to a “bending of the rules” in this example (*ibid.* 392).

A more problematic argument used by Allen to dismiss examples such as (88) and (89) is that the initial elements are full NPs and, because case marking on nominals is lost, it cannot be said with certainty that they are nominative subjects and not dative-fronted passives (1995: 391).

- (88) *The Duke Mylon was **geven** hys liff, and fleygh out of land with hys wyff*
 the Duke Mylon was given his life, and fled out of land with his wife
 ‘Duke Mylon was given his life, and fled out of the country with his wife’
 (Ric.Coer de L. (Auchinleck) 1307; Allen 1995: 390, her 249)
- (89) *Ak Ihesus nys on eorþe nouȝt so muche goed **byleued**...*
 but Jesus not-is on earth not so much good left ...
 ‘But Jesus is not left so much good on earth..’
 (SP 1607; Allen 1995: 392, her 254)

These verbs and types of passives still only occurred with fronted dative pronominal recipients, as the following example shows.

- (90) *þe nas so muche goed **ygraunted***
 Thee not-was so much good granted
 ‘So much good was not granted to thee.’
 (SP 1611; Allen 1995: 392, her 255)

Because of the existence of passives with a pronominal dative recipient, Allen concludes that it cannot be said that the nominal recipients in (88) and (89) are nominative subjects or fronted datives. While this may be true, it does not rule out the possibility that they are nominative subjects, as Van Kemenade (1998) points out. The relevant argument is not that they are definitely not recipient passives or definitely not dative-fronted passives, but rather that it cannot be stated with certainty that they are either one or the other. This ambiguity is an interesting point

in itself, because it could mean the status was ambiguous for speakers as well, possibly already paving the way for the recipient passive.

The earliest convincing examples, according to Allen, date from the late fourteenth century, and are given in (91) and (92) (these are from Visser's original collection).

- (91) *Item as for the Parke she is a lowyd Every yere a dere*
 'Item: as for the park, she is allowed a deer each year'
 1375 (AwardBlount p.205; Allen 1995: 393, her 256)

- (92) *playnly bu are forbodyn bobe*
 plainly thou-NOM art forbidden both
 'Plainly, you are forbidden both.'
 (Wyc.Wks XXVI 383.24; Allen 1995: 393, her 256)

Allen states there are more examples in texts written by Wycliffe, and that nominal subjects in his writing can also be analysed as recipient passives because no dative-fronted passives occur in these texts, only PP-fronted passives, showing that fronting of indirect passives was no longer possible (1995: 393).

Some additional early examples are given by Denison (1993).

- (93) *and bey shall be assigned redy shippyng and passage*
 And they shall be allotted ready shipping and passage
 1418 (*Let. War France* in *Bk.Lond.E.* III.xvi 77.8)
- (94) *Now is he nayled to the cros and youen eyssel medled with galle to drynke*
 now is he nailed to the cross and given vinegar mixed with gall to drink
 a1450 (*Aelred Inst.*(2) (Bod) 846)
- (95) *if Ser Thomas thynk that he shuld be a-lowyd mo, he shall be*
 if Sir Thomas think that he should be allowed more he shall be
 'if Sir Thomas thinks that he should be allowed more, he shall be'
 c1453[not before] (*Paston* 47.9)
- (96) *and whan he was gyvyn the gre be my lorde kynge Arthure*
 and when he(SUBJ) was given the prize-for-victory by my lord King Arthur
 'and when he was given the prize by my lord, King Arthur'
 a1470 (Malory, *Wks.* 699.19)
- (97) *tyll tyme pat I be paied fully my salary*
 Till time that I be paid fully my reward
 'till the time that I am paid my reward in full'
 a1500(?a1450) (*Grom.* (Hrl) 88.18) (Denison 1993: 111, his 46-50)

Note that in example (94), in contrast to (90), it is clear that the ellipited subject is a nominative because of the nominative subject *he* in the first part of the sentence. Denison gives 5 early examples, while Allen does not list all her examples nor does she state exactly how many she has found in total in the fourteenth and fifteenth century, which means it is not possible to give a total number of examples in this period. The verbs in the examples given by Denison and Allen are *allow* (2), *forbid*, *assign*, *give* (2), and *pay*.

In conclusion, the first examples show up around 1400 and there are a handful of examples in the fourteenth century. This is about half a century later than the prepositional passive and a considerably slower introduction, with still only a small number of examples around 1400.

3.3.2.2 Previous accounts

Across the literature, there seem to be two types of explanation for the introduction of the recipient passive: one which focuses on the loss of the distinction between direct and indirect objects, and another which focuses on the reanalysis of any preverbal element as a subject. Jespersen (1909-49: III 299-311), as discussed by Denison (1993: 113-114), combines the two approaches, proposing that the loss of the dative/accusative distinction in the active led to a situation where any object, direct or indirect, became available for passivization. On the other hand, the loss of case marking also led to a loss of distinction between subject and non-subject, leading to a fronted dative to be reanalysed as a subject: *The girl (DAT) was offered a watch* becomes *the girl (NOM) was offered a watch*.

Lightfoot (1979: 259-265) connects the introduction of the recipient passive to an overall change in the nature of the passive: the addition of a transformational passive to a lexical passive, a comparable distinction to verbal vs adjectival passives (cf. Section 2.3.1). The transformational passive, crucially, could target arguments other than direct objects. In a later publication, Lightfoot (1981) seems to abandon this account and instead attributes a central role to structural and oblique Case assignment and the loss of Case marking, later followed by Van Kemenade (1987: 211). In this account, the recipient passive is impossible in OE because the indirect object receives oblique (dative) case in its base-generated position; it cannot move to the subject position because it would then be assigned another Case, resulting in a “case-clash” (1987: 211), under the assumption that structural (nominative) Case is assigned to the subject position. Lightfoot (1981) and Van Kemenade (1987) claim that oblique case was lost in (late) ME—in turn due probably to the loss of Case marking (1981: 107)—and that the “case-clash” problem had now been resolved, making the recipient passive grammatical. Again, then, this is an account which builds on the blurring of distinctions between direct and indirect objects, and one which accounts for the change by an extension of the types of objects that could passivize. A problem with the the loss of case marking as the most important factor (although it may be a necessary condition) is the observation that Dutch also lost case marking but did not acquire a recipient passive.

Other authors have focused solely on the development of the subject as the key to the introduction of the recipient passive. Van der Gaaf (1929: 60-62) proposes that the nominal recipient (the “personal object”) came to be reanalysed or interpreted as the subject, because it occurred in the position “usually taken by” the subject (61). The dative was lost, Van der Gaaf claims, in the “confusion” about case marking in the ME period (1929: 1). Visser (1973: §1967) and Mitchell (1985: 360) present a similar scenario, saying the development is analogous to the change of ‘me thinks’ to ‘I think’ (Visser 1973: 2144). In other words, there is supposedly a relation between the loss of dative subjects in sentences like (98), which were replaced with (99), and the loss of dative-fronted passive and the rise of the recipient passive.

(98) Him was pleased.

(99) He was pleased.

(Allen 1995: 349)

The change is that the dative object which formerly retained its dative case when it was passivized, as in (117), was now used with nominative case in the passive. According to Allen (1995: 350), the passive in (99) had replaced the passive in (98) by the beginning of the thirteenth century. Allen points out that this replacement does not represent a grammatical change, but only a change in Case marking (*ibid.* 377). This claim is especially interesting because the change seems similar to the change from *Him was given the book* to *He was given the book*.

Allen (1995) discounts the scenario which attributes the introduction of the recipient passive to the loss of case marking, as well as the replacement of the dative-fronted passive *Him was given the book* with *He was given the book*. In both cases, her main argument is the timing of the developments. She states that the change from (98) to (99) took place much earlier than the introduction of the recipient passive: after this development in the thirteenth century, it takes another century and a half before the first recipient passives are attested. Moreover, the dative-fronted passive, which the recipient passive presumably replaces, was no longer in use by that time, as she states in the following passage:

[T]he dative-fronted passive became infrequent in prose by the beginning of the thirteenth century, and becomes quite unusual by the middle of the fourteenth century, even in the rather archaic *Azenbite of Inwyt*. It is questionable whether the dative-fronted passive was part of ordinary speech by the late fourteenth century, when the recipient passive was introduced, and so could have been the basis of a reanalysis. (Allen 1995: 380)

The loss of case marking, too, precedes the introduction of the recipient passive: Allen shows that the case markings on nouns were lost long before the introduction of the recipient passive (*ibid.* 415). This conclusion by Allen relies heavily on the gap between several developments, but ‘the middle of the fourteenth century’ for the

though it complicates deducing the semantic relations (425). This means that in (100)a, the NP_{REC} would now be analysed as a direct object and passivized accordingly.

Allen's approach relies heavily on the presence of word orders in a limited corpus, which could be problematic as evidence. Her account does provide a more detailed version of the conflation of direct and indirect objects and finds a specific reason in the language of the relevant period, something which not all other accounts that assume this conflation, like for instance Lightfoot (1981), do in such detail. However, the connection she makes to other changes in this period fails in one important aspect: although it is clear that the loss of dative-fronted passives due to the loss of dative fronting in general is a consequence of the loss of V2, Allen does not mention this important syntactic change that cannot be excluded from any treatment of ME syntax.

Allen's account also relies heavily on dismissing the suggestion that the preverbal argument was reanalysed as the subject in dative-fronted passives, but we have seen that this gap between loss of the dative-fronted passive and the introduction of the recipient passive might not be as wide as she claims. The gap between loss of dative subjects and the introduction of recipient passives, which she also uses as an argument, could in fact be a strong factor in favour of the introduction of the recipient passive: in the total absence of dative subjects of impersonals with the additional later loss of dative fronting in all contexts, preverbal NPs can only be analysed as nominative subjects. We will turn to the available data in the PPCME2 to examine this question below, but first, I outline a scenario which connects the loss of the dative-fronted passive to the introduction of the recipient passive, but in a different way from the scenarios discussed in this section.

3.3.3 Revisiting the importance of the loss of the dative-fronted passive

There is a central theme running through these accounts for the introduction of the recipient passive, namely reanalysis or reinterpretation either of the pre- or the postverbal argument. Almost all authors refer to the fact that case distinctions between nominative, dative and accusative were lost for nominal arguments, and that SVO word order became fixed. With the loss of case marking, speakers may have relied increasingly on word order cues rather than case marking cues, as Allen (1995) argues in her account. The question for the recipient passive, then, seems to be whether it is the case that the clause-initial NP is interpreted as a subject rather than a dative now that case is no longer marked, or whether it is the case that the dative/indirect object is reanalysed as an accusative/direct object, and can accordingly be passivized following standard passivization processes. In this section, I argue that both factors are relevant, in that both the changes in the preverbal position (loss of dative subjects) and postverbal position (changes in order of objects) play a role in allowing the recipient passive. However, I propose that the introduction itself is due to the loss of the dative-fronted passive.

3.3.3.1 Dative-fronted passive blocks the recipient passive

The analysis presented in Section 2.3.3 where the passive targets the nearest available NP to become the subject, regardless of whether it is an indirect or direct object, builds on different underlying structures and does not rely on semantic roles or (structural or oblique) case. The basic analysis of the recipient, theme and dative-fronted passive has already been outlined in Section 3.3.1.2. Let us turn to the passives that are attested in OE. Most importantly, before the introduction of recipient passive, OE has both the theme passive (with two NP objects), as in (101), and the dative-fronted passive, as in (102).

(101) The book.NOM was given Mary.DAT.

(102) Mary.DAT was given the book.NOM.

As previously explained, I will assume that these passives are both based a similar original (i.e. merged) structure—*gave the book.NOM Mary.DAT*—with subsequent passivization of *the book*. The difference between the two passives is that in the dative-fronted passive, the recipient dative *Mary* is fronted, or topicalized, to the Left Periphery because of V2, and the passivized subject remains low.³⁹ In other words, the dative is not passivized but only fronted. The PP-fronted passive—*To Mary was given the book*—which becomes possible as PP recipients become more frequent has a similar analysis to that of the dative-fronted passive: it is essentially a theme passive, but with a fronted PP. The fronting of the PP is determined by general fronting constraints.

The recipient passive is crucially different because it is indeed the recipient that is passivized, but from a different original structure, namely *gave Mary the book*.

(103) Mary.NOM was given the book.

I would like to propose that as long as the dative-fronted passive is possible in English, the recipient passive, i.e. with nominative subject, is ungrammatical because there is already a way to place the recipient in first position, with a similar effect to passivization, namely the dative-fronted (theme) passive. Only when the dative-fronted passive is lost, a development contingent on the loss of V2, does the recipient passive, with a nominal subject recipient, become possible.

The same system that is in place in English up to the loss of the dative-fronted passive still seems to be in place in the modern V2-languages Dutch and German,

³⁹ With the dative in topicalized position (Spec,CP or Spec,FP), the finite verb moves to the Left Periphery as well (to F). Of course, in many dative-fronted passives, the passivized argument occurs after the passive participle and not, as would be expected with movement of the finite verb across the subject position, between the finite verb and the subject. The details of this analysis go beyond the scope of this chapter, but see Sections 2.2.1 and 4.2.2.4 for an account of this and observations on the possibility for passive subjects to remain in a low position within the VP.

which have a theme passive as well as a dative-fronted passive. The following examples illustrate the theme passive (104) and dative-fronted passive (105) in Dutch.

(104) *Het boek.NOM werd Marie.DAT gegeven.*
The book was Mary given

(105) *Marie.DAT werd het boek.NOM gegeven.*
Mary was the book given

The situation in Dutch is different from OE, however, in that the standard order of elements in actives with two NPs is recipient-theme, as illustrated in (106), although Dutch allows for different orders when the arguments are pronominal, witness (107).

(106) *Jan gaf Marie het boek.*
Jan gave Marie the book

(107) *Jan gaf het haar.*
Jan gave it her

The analysis presented earlier would predict that the recipient passive *Marie.NOM was given the book* would be grammatical and not the theme passive *The book was given Mary*, but this is clearly not the case for Dutch. This means we need to account for the grammaticality of (108) and ungrammaticality of (109).

(108) *(Aan) haar werd het boek gegeven.*
(To) her was the book given

(109) **Zij werd het gegeven.*
She was it given

There are two possibilities to explain the unacceptability of (109). One is to say that there is in fact a principle in Dutch that only direct objects can be targeted by passivization. Although case marking is lost in Dutch, it could be that Dutch speakers still have more clues that the immediately postverbal NP has dative Case, although it is not immediately clear what these clues would be once case markings on nominals are lost. A more attractive option, in line with the proposal for the introduction of the recipient passive in English, is to say that because of the existence of the dative-fronted passive, (109) is out: either nominative recipients are not allowed or nominative recipients are not necessary in the language because there is already a way to front recipients. This would be a similar situation as in ME: the dative-fronted passive blocks the recipient passive.

In this scenario, the difference between OE on the one hand and Dutch and German on the other hand, is that dative/indirect object fronting was never lost in Dutch and German. Dutch and German still have V2, which means the dative-fronted passive remains in use and blocks the formation of a true recipient passive *Zij werd het boek gegeven* from *Hij gaf haar het boek*. The loss of dative fronting in English, related to the loss of V2, sets English apart from Dutch and German, and accounts for the different development in English as the consequence of other changes which took place in English in the period that the recipient passive became possible. The following section presents evidence for this scenario from a study of the PPCME2.

3.3.3.2 Evidence from the PPCME2

The corpus study yields three observations that support the scenario outlined above: (i) the introduction of the recipient passive can be dated to the late fourteenth century; (ii) there are indications for a later date for the loss of the dative-fronted than previously suggested; and (iii) the type *They gave the book Mary*—the sentence type that was also crucial to Allen’s scenario—disappears before the passive version *The book was given Mary* disappears.

Recipient passives: no earlier examples than provided by Allen (1995)

The PPCME2 coding contains two different labels for objects, crucially making a distinction between ‘first’ objects, roughly corresponding to direct objects (labeled NP-OB1) and ‘second’ objects, roughly corresponding to indirect objects (NP-OB2). The distinction is not entirely straightforward because additional material is included in both groups (predicate nominals for NP-OB1 and “other oblique uses” for NP-OB2, see Kroch & Taylor 2000). The same labels are used in the passive versions of these ditransitives to represent role that the objects would have had in the active. This makes it possible to select recipient passives with the following query.

Query 3.2 Recipient passives	
e.	IP* idoms *SBJ*
f.	IP* idoms *BED* *BEP* BE BAG BEN
g.	IP* idoms *VAN*
h.	IP* idoms NP-OB1

The query first selects clauses with a subject and a passive: a form of *to be* and a passive participle, both as main constituents of the IP (line 3.2a-c). It furthermore stipulates that these passives also dominate a direct object (line 3.2d). The indirect object has been passivized to become the subject of the passive and is no longer represented in the passive. However, not all examples that result from this query are in fact recipient passives, so they needed to be sorted by hand.

The occurrences of recipient passives in the corpus confirm Allen’s findings that the first recipient passives only start to appear in the late fourteenth century.

There is one suggestive example from the middle of the fourteenth century, but there are a number of problems with this sentence, given in (110).

- (110) *Ich am geuen in langour,*
I am given in suffering.
'I am handed over to suffering'
c1350 (cmearlps,107.4680)

First of all, this instance is given as the only example of *given in languor* in the MED and it is peculiar because it does not seem to be the passive of a relevant (i.e. clearly ditransitive with two NPs) use of *give*. Secondly, later translations of this psalm have quite a different translation—*I am shut up* (King James, Psalms 88:8) and *I am taken in* (Wycliffe)—which seems to indicate that the speaker of this sentence is himself the object that is being given, instead of the recipient of some gift (cf. Russom 1982: 679 for a similar argument about another early example of a supposed recipient passive presented by Lieber 1979). Finally, the same text also has an example of a dative-fronted passive, given in (111).

- (111) *and hym shal bi geuen of þe seruice of gode;*
and him shall be given of the service of God
c1350 (cmearlps,86.3761)

The fact that dative-fronted passives are still found in the same text makes it seem unlikely that *Ich* in (110) is unambiguously a nominative recipient.

The following examples represent all recipient passives found in the PPCME2.

- (112) *for þe prioresse is geuin a mater to be proud in þe begynnyng of hyr ordinance, when sho is geuyn in mynde be hir awne thoghtis þat sho is lowsid owte of þe power of hir abbesse,*
for the prioress is given a matter to be proud in the beginning of her ordinance, when she is given in mind by her own thoughts that she is freed out of the power of her abbess
a1425 (cmbenrul,43.1346)
- (113) *And yef þu gruches & dos þat ilke þing, god es noht paid þar-of;*
and if you complain and do that same thing, God is nothing paid thereof
a1425 (cmbenrul,10.341)
- (114) *and þei þat took hem were treuly payed too þousand pound.*
and they that took him were truly paid two thousand pound
a1464 (cmcapchr,153.3585)

- (115) *and they were servyd nexte unto the quene every cours coveryde as the quene;*
 and they were served next to the queen every course covered as the queen
 c1475 (cmgregor,139.584)

Note that example (113) has a nominal subject, but the example below from the same text shows that *payde* already occurs with pronominal subjects (although there is no direct object in this example).

- (116) *he giue vs for his grace þat we may sua wirke and sua selle, þat he be payde þar-of.*
 he gives us for his grace that we may so work and so sell, that he be paid thereof
 a1425 (cmbenrul,37.1207)

The fact that only a small number of examples are found in the PPCME2 corresponds to Allen's (1995) findings. These low numbers indicate that the introduction of the recipient passive was slower than the introduction of the prepositional passive—of which more examples occur in the PPCME2 (cf. Section 3.2.2.2)—but at the same time, it can be seen as the logical consequence of the greater lexical restrictions on the recipient passive (i.e. the number of available ditransitive verbs) and the fact that the theme passive with two nominal NPs is still possible at this stage.

Dative-fronted passive: later examples than provided by Allen (1995)

Allen (1995) dismisses the scenario where the loss of the dative-fronted passive led to the introduction of the recipient passive, based on her observation that the dative-fronted passive was no longer in common use in the middle of the fourteenth century (1995: 380). The PPCME2, however, gives later examples of dative-fronted passives. The following query was used to select the dative-fronted passives in the PPCME2.

Query 3.3 Dative-fronted passives
a. IP* idoms *BED* *BEP*
b. IP* idoms *VAN *
c. IP* idoms NP-SBJ*
d. IP* idoms NP-OB*
e. NP-OB* precedes NP-SBJ*

The query selects clauses with a subject, a finite form of *to be* and a passive participle (line 3.3a-c). It then selects clauses which have an object, which precedes the subject (line 3.3d-e). I included both first and second objects in this study in order to make sure that I did not miss any examples because of errors in the coding.

Contrary to Allen's (1995) claim, I have found instances of dative-fronted passives later than the middle of the fourteenth century. Examples (117)-(120) show the last examples of dative-fronted passives that I found in the corpus.

- (117) *þo hym was **i-graunted** riȝt in þe electioun of þe pope*
then he was granted right in the election of the pope
a1387 (cmpolych,VI,267.1945)
- (118) *þefore þe kyng Gurthrym, þat we clepeþ Gurmundus, **were i-geve** þe*
provinces of Est Angles and of Norphomberland for to wonye ynne.
therefore the king Gurthrym, who we call Gurmundus, was given the
provinces of the East Angles and of Northumberland to live in
a1387 (cmpolych,VI,377.2769)
- (119) *ne hym **was owed** no sych rewme, sib God wolde not þat hit were so.*
not him was owed no such room, since God would not that it were so
c1400 (cmwycser,I,415.3413)
- (120) *Yef hir **be sente** ani þing fra hir frende, sho sal noht recaue it,*
if her be sent any thing from her friend, she shall not receive it,
a1425 (cmbenrul,36.1159)

The last two examples, from the *Wycliffite sermons* and the *Northern Prose Rule of St. Benet*, are especially interesting, because (112) above showed that in the *Rule of St. Benet* recipient passives were already possible, while example (121) shows that in the *Wycliffite sermons* fronting of the PP in the passive is also possible. While the example does not have two NP arguments, but one clausal argument, and the sentence does not have exactly the same structure as a dative-fronted passive, the example does clearly show the fronting of a PP recipient.

- (121) *And Crist seyde vnto hem þat to hem was **grawntid** to knowe þe pryuyte of þe*
rewme of God
and Christ said to them that to them was granted to know the privacy of the
realm of God
c1400 (cmwycser,385.2845)

The dative-fronted examples in this section show that the dative-fronted passive had not completely disappeared from texts at the time that the first recipient passives are attested, i.e. the late fourteenth century. They also show that in some cases both the recipient passive and the dative-fronted passive occur in one text (*The Rule of St. Benet*). This means that the gap between the disappearance of the dative-fronted passive and the introduction of the recipient passive can no longer be taken as an argument against a relation between the two developments, as argued by Allen (1995).

The disappearance of *He gave the book Mary* and *The book was given Mary*

One final issue is the relevance of the order *He gave the book Mary*, which Allen (1995) takes to be crucial to the introduction of the recipient passive, because in its absence *Mary* in *He gave Mary the book* would now be analysed as a direct object. Moreover, the proposed close relation between active and passive in the analysis presented in Section 3.3.1.2 raises the question whether it is true that the loss of the active version predates the loss of the passive version, *The book was given Mary*. The loss of this order means that the corresponding theme passive *The book was given Mary* would also start to disappear.

Let us first consider the order theme-recipient, as in *He gave the book Mary*. By the M3 sub-period (1350-1420), this order occurs almost solely with pronominal arguments.⁴⁰ The only two examples in this period that do not have two pronominal arguments are given in (122) and (123).

(122) *but opere kynges þat were of more strengþ þan he, binome him his lande.*
 but other kings who were of more strength than he, took-from him his land
 'but other kings, who were stronger than him, took his land from him'
 c1400 (cmbrut3,23.684)

(123) *for they bireven alle thise hym that sholde worshipe Crist and hooly chirche,*
and preye for Cristene soules.
 for they take-away all this from-him that should worship Christ and holy
 church, and pray for Christian souls
 'For they take away everything from him who worships Christ and the holy
 church, and prays for Christian souls'
 c1390 (cmctpars,320.C1.1366)

The first example has one pronominal element, *him*, and one nominal argument, *his lande*. The second example has an NP, *alle thise*, and a postmodified pronoun as the second element, *hym that sholde worshipe Crist and hooly chirche*. These two examples do not have two straightforward NP arguments, and it is clear that the pronominal version is already the preferred possibility in this period.

As for the corresponding passive *The book was given Mary*, the following examples show the last examples that occur in the corpus.

(124) *and anone hit was tolde þe Quene how þat þe Kyng was comen allone wipout*
company;
 and immediately it was told the Queen how that the King was come alone
 without company

⁴⁰ The relevant labels that were used to select these clauses are: NP-SBJ*, finite verb, NP-OB1 and NP-OB2. Pronominal objects were selected using an additional requirement that the object only dominates PRO*.

‘And immediately the Queen was told that the King had come alone, without company;’

c1400 (cmbrut3,117.3546)

(125) *Whanne þat was i-tolde þe kyng, þe kyng was swiþe wrooth,*

‘When that was told the king, the king was very angry’

a1387 (cmpolych,VI,71.499)

(126) *Þis was i-tolde þe bisshop,*

‘This was told the bishop’

a1387 (cmpolych,VIII,91.3602)

Note that all subjects in these examples are pronominal, perhaps suggesting that they are passives of the version *He gave it Mary* rather than *He gave the book Mary*. These examples are all dated in the M3 (1350-1420) period. Although the PPCME2 provides only a handful of examples of the relevant sentences, this seems to suggest that the theme passive of two NPs indeed starts to disappear after the corresponding active is lost.

3.3.3.3 Conclusion

In this section, I outlined a scenario where the dative-fronted passive was the crucial factor for the introduction of the recipient passive. I showed that the PPCME2 contains examples of dative-fronted passives dated at the same time that the recipient passive was already attested (around 1400). I proposed that the existence of the dative-fronted passive blocked the recipient passive, although there may be a time of overlap where both nominative and dative recipients in initial position are possible. When we consider other West-Germanic languages, it is clear that Dutch and German do have the two actives but only allow the theme passive. However, both these languages have the dative-fronted passive available for the sentence *John gave her the book*, and the difference must then lie in the fact that they retain fronting of dative objects, making the true recipient passive ungrammatical.

3.3.4 Summary and conclusions

This section set out to establish the time of introduction of the recipient passive, as well as the reason for its introduction, based on a re-examination of the literature and a corpus study of the PPCME2. The corpus study confirmed earlier conclusions that the recipient passive is introduced in the late fourteenth century. It provided, on the other hand, counterevidence for the claim that the dative-fronted passive was lost well before the recipient passive was introduced, with examples from around 1400. I argued that the loss of the dative-fronted passive was the crucial factor for the introduction of the recipient passive, because as long as fronted dative

recipients were possible the recipient passive with nominative recipient subjects was blocked.

The introduction of the recipient passive seems to be much slower than the introduction of the prepositional passive. After the introduction in the late fourteenth century, there are only a handful of examples in the PPCME2, corresponding to earlier findings by Allen (1995) and Denison (1993). This is no surprise considering the small number of ditransitive verbs and the competition with the active ditransitives with an NP and a PP argument and their corresponding passives. The data from the PPCME2 furthermore showed that the dative-fronted passive remained possible in some texts until roughly 1400, leading to a situation where both the dative-fronted passive and the recipient passive are attested in the same period, in contrast to Allen's findings and indeed, her scenario for the introduction of the recipient passive.

To explain the introduction of the recipient passive, I proposed a scenario in which the dative-fronted passive, which I analysed as a theme passive with a fronted dative, blocked the occurrence of nominative recipient arguments in preverbal position. With the loss of dative fronting, likely to be related to the bundle of changes that make up the loss of V2, recipient passives became possible. In other words, the crucial factor in the introduction of the recipient passive is the loss of fronting possibilities for datives, in turn also related to the factors that have previously been said to be the causes for the introduction, the loss of (inherent) case marking (cf. Lightfoot 1981), loss of distinctions between dative and accusative objects (cf. Allen 1995) and the fixing of a strict SVX order (cf. Van der Gaaf 1929).

The scenario proposed explains the difference with German and Dutch, closely-related languages which do not have a recipient passive. These languages, which in contrast to English continued to show the V2 constraint, never lost the dative-fronted passive and so were never in a position to develop a recipient passive. Chapter 5 addresses the question of the development of the recipient passive in the periods after ME.

Section 3.4 The introduction of the ECM passive

This section investigates the introduction and development of the so-called Exceptional Case Marking passive, another passive that was introduced in the English language around 1400 AD, at roughly the time that the V2 constraint was starting to decline. The introduction of the ECM passive is different from that of the recipient and prepositional passives because it is not just the passive form that is introduced, but the entire ECM structure. This means that the development of the active is just as important as the development of the passive, something which becomes even more important when we consider that the ECM throughout its history occurs more frequently in the passive than in the active, in sharp contrast to general active-to-passive ratios. In this section, I argue that the passive version caught on in English more easily than the active because it is structurally very similar to passives that were already present in English, meaning that—in contrast to the

active ECM—it was not a marked construction. Crucially, I argue, there is no evidence of a direct relation to the loss of V2, both structurally and in terms of timing.

The introduction of the ECM construction, shown in (127) in its active form and in (128) in its passive form, has generally been dated to the late fourteenth century (Warner 1982, Fischer 1989).

(127) They believed John to be a liar.

(128) John was believed to be a liar.

In comparison to the prepositional and recipient passives, there has been less discussion about what constitutes the earliest unambiguous example of the ECM. As was the case for those passives, however, the question of why the construction was introduced has received much attention. One important factor proposed to contribute to the introduction is the influence of Latin (Warner 1982), based on the observation that many early ECM examples occur in translations from Latin. Yet this Latin influence has been doubted as the only factor by Fischer (1989) and Los (2005), who propose another source for the development of ECM: reanalysis of an already existing object control construction. Another proposed factor leading to the introduction and establishment of the ECM passive is the the loss of V2, as outlined by Los (2005); it is specifically the ECM passive that Los first proposed to be introduced as a consequence of the loss of V2. This calls for a re-examination of the data in light of this hypothesis, taking into account, once again, both the time of the introduction with respect to the loss of V2, as well as factors possibly connected to V2 leading to the introduction of the ECM construction and specifically the passive.

One aspect which has puzzled researchers for years is the observation that the ECM is much more frequent in the passive than it is in the active, in contrast to regular distribution of actives and passives. Interestingly, this dominance of the passive has been found to exist not only in PDE (Mair 1990, Noël 2008) but already from the earliest occurrences in ME onwards (Warner 1982, Fischer 1994). For the first examples, Warner (1982: 136) already notes that the passive is less dependent on the contexts of Latin translations than the active. The dominance of passive over active in PDE has in turn led Noël (2008) to propose that the passive has been grammaticalized into a fixed template with a specific function of encoding ‘evidentiality’. Others have referred to the discourse function of the passive, which the active version lacks (Mair 1990, Los 2005). Los’s (2005) proposal takes this unbalanced distribution into account, with a focus on the introduction of the passive in particular, but as we will see, there are some problems with this scenario, most importantly the timing of the introduction with respect to V2.

This section is structured as follows. Section 3.4.1 discusses the structure and use of the ECM construction, both active and passive, in PDE, with a syntactic analysis of active and passive ECM in Section 3.4.1.2. Section 3.4.2 discusses the previous accounts for the introduction of the ECM, reviewing the influence of Latin (Warner 1982) and the reanalysis of object control constructions (Fischer 1989, Los

2005). Section 3.4.3 then presents the results of the corpus study of the ECM in the PPCME2. This is used as the basis for an account of the dominance of the passive with ECM verbs based on similar passives that were already present in English. Section 3.4.4 concludes.

3.4.1 A characterization of the ECM passive in Present-day English

The ECM in PDE has not only been considered exceptional in terms of its syntactic structure, i.e. its ‘exceptional’ configuration of case marking, but it has also been noted for its unusual distribution of active and passive occurrences, with the passives being much more common and used with a wider range of verbs (Mair 1990, Noël 1998). This has led researchers to focus on the properties of the passive separately, with Noël (2008) arguing that it has grammaticalized, in the sense that the form has been linked to one specific function, while Mair (1990) and Los (2005) focus on the discourse function of the passive subject.

3.4.1.1 The ECM in Present-day English

The ECM construction—sometimes referred to as accusative-cum-infinitive (a.c.i) after its Latin model (e.g. Fischer 1989), Raising to Object (e.g. Postal 1974) or the descriptive VOSI (Verb-Object/Subject-Infinitive, Visser 1973) or SVOC (Subject-Verb-Object-(Predicative) Complement, Mair 1990)—contains, in the active as in (129), a matrix verb followed by an NP and (subjectless) infinitival clause. In the passive, as in (130), this object of the main clause/subject of the subordinate clause has been passivized to become the subject of the main clause.

(129) They believe him to be a liar.

(130) He is believed to be a liar.

The structure in (129) looks similar to object control clauses as in (131), in which the main clause verb is also followed by an NP and an infinitive clause.

(131) They commanded him to leave the room.

However, there is a difference in the status of the NP in these two types of clauses. Object control clauses occur with so-called three-place verbs, which take as their complements a subject, an object and an infinitive clause. The NP in (131), then, is an argument of the verb, and there is an empty subject (PRO) in the infinitive clause. ECM clauses, however, occur with verbs that normally take just two arguments, such as *believe*—these verbs are referred to as monotransitive or two-place verbs, i.e. verbs that have two arguments: a subject and an object. This means that in (129) above, *him to be a liar* is one argument of the verb; in other words, the NP *him* is not itself an argument of the verb *believe*. The crucial difference between object control and ECM lies in the fact that the NP in the *NP to VP* sequence in object control really is the object of the higher verb, while in ECM it is, at least thematically,

an argument of the lower clause, possibly raised to become an argument of the main clause (see Postal 1974). This raising explains its accusative case and gives a structural account of the observation that the relevant NP is thematically both the object of the main clause and subject of the subordinate clause.

The distinction between two-place and three-place verbs is well-accepted in the literature, although there has been much discussion on which verbs belong to which group. The groups are usually called after clear-cut members, like *believe* for two-place and *want* for three-place verbs (Postal 1974) or *expect* for two-place and *persuade* for three-place verbs (Fischer 1989). Postal (1974: 298) claims that the group of unambiguous ECM verbs, i.e. two-place verbs but with seemingly three arguments, is usually taken to consist only of *believe*, *prove*, *find* and *show*, but adds quite a number of verbs, presenting the following list of verbs, which Los (2005: 254) later calls “verbs of propositional attitude”.

(132) *Acknowledge, admit, affirm, allege, assume, believe, certify, concede, consider, declare, decree, deduce, demonstrate, determine, discern, disclose, establish, estimate, feel, figure, find, gather, grant, guarantee, guess, hold, imagine, intuit, judge, know, note, posit, presume, proclaim, prove, reckon, recognize, remember, report, reveal, rumour, say, show, specify, state, stipulate, suppose, surmise, take, think, understand, verify.*

(Los 2005: 254, based on Postal 1974: 305-310)

Fischer’s (1989: 174-175, 184) *expect* group for OE includes several subgroups distinguished by Visser (1973: 2250ff), such as perception verbs (e.g. *behold, feel, find, hear, see*), causative verbs (e.g. *let, make, cause*), verba declarandi et cogitandi (‘verbs of thinking and declaring’) and the verbs expressing a wish or desire. Note, however, that it is only the last two groups, which includes verbs such as *wish, like, think, consider, declare*, that are relevant to the development of the *believe*-ECM in ME: it is these verbs that start to occur with the *to*-infinitive in an ECM configuration.

One complicating factor in classifying the verbs is the observation that some verbs of the *want* group occur with both ECM and object control. Postal (1974: 318-321) identifies these as *allow, permit* and *order*. Examples (133) and (134) from Postal illustrate.

(133) I allowed [Bob to leave] ‘I gave permission for Bob to leave’

(134) I allowed Bob [PRO to leave] ‘I gave Bob permission to leave’

In the first interpretation, Postal states, raising has taken place, i.e. Bob originates in the lower clause but moves to the main clause, whereas in the second interpretation, Bob originates as the object in the main clause. Postal discusses numerous tests which can be used to determine whether raising has taken place, i.e. whether the NP becomes the object of the higher clause. One such test is to

determine whether there is a semantic contrast between the active version and a version in which the infinitive is passivized.

- (135) a. I convinced Bob to interrogate the witness
b. I convinced the witness to be interrogated by Bob.

- (136) a. I expect Bob to interrogate the witness
b. I expect the witness to be interrogated by Bob

If there is a semantic contrast, as in (135), the NP is truly an object of the main clause three-place verb, while if there is no contrast, as in (136), the entire infinitive clause is the complement of the matrix verb. Another test to determine the status of the ambiguous argument that is discussed in the literature, and also specifically for ECM in the history of English (e.g. Postal 1974: 151, 320; Mair 1990: 166; Los 1999: 294-5; Fischer 1989: 176-177), is animacy, or ‘mind-possessing’ in Postal’s terms: when the argument is inanimate, it is likely to be an ECM construction. When the argument is animate, both an ECM (two-place) and an OC (three-place) interpretation are possible, as we saw in examples (133) and (134).

Several authors have noted that the passive ECM is more common than the active ECM in PDE. Postal (1974: 298-301) already notes the restriction on ECM in the active and in fact, his main reason for extending the group of genuine ECM verbs beyond *believe*, *prove*, *find* and *show* is that there are many verbs that allow this pattern in the passive even if they do not allow it in the active. Mair (1990) presents numbers for the frequency of verbs in several infinitival complement patterns in the Survey of English Usage, including the ECM configuration—his SVOC pattern (Subject-Verb-Object-Complement). He finds that passives in this pattern are roughly twice as common as the actives (64 actives against 109), and calls the passive the “statistical norm” (1990: 175-176).⁴¹ He also repeats the especially striking observation earlier made by Postal that some verbs do not occur in the active at all. In Mair’s classification, ECM verbs can also be classed as a monotransitive pattern with a *to*-infinitive which has a subject, such as *expect*. He states that a part of these—the relevant ECM verbs—occur “freely, frequently or even predominantly” in the passive (1990: 113).⁴² The ‘skewed’ distribution of active and passive is further confirmed by Noël’s (2001) study of these verbs in the BNC: some of the most common verbs in the relevant pattern, *say*, *repute* and *rumour* (2001: 257-259, Table 1) only occur in the passive and not in the active, while all of the twenty most frequent verbs overall occur more often in the passive than the active. Only among less frequently used verbs do verbs occur with a higher frequency for the active than the passive.

⁴¹ Mair excludes 63 examples of *be supposed to* on the ground that this cannot be seen as a grammatical passive of a productive active, because it has acquired a separate, additional meaning.

⁴² Interestingly, he also finds that clauses taking an object and an infinitive (so object control verbs) also occur more frequently as passives than the general rate of active and passive: these have a roughly 3:1 active-passive ratio, against the regular 9:1 found in Svartvik 1996 (Mair 1990: 158-159).

In the synchronic literature the dominance of passive over active has been given various explanations. Noël (2001) proposes that the passive version of the ECM is a grammaticalized construction which has acquired an ‘evidential’ reading; a function he explains as “the expression of the kind of evidence a person has for making factual claims” (2001: 266, based on Anderson 1986: 273). Indeed, many instances of ECM passives can be interpreted in such a way, a type of use especially found in news reporting. Consider the following example.

- (137) Both **are said to be** of Chechen origin, and **are reported to have moved** to the United States about 10 years ago.
(www.bbc.com, 19 April 2013)

This sentence, which contains two ECM passives, occurs in a news article published on a developing news story (the manhunt on two suspects of the Boston Marathon bombings on 15 April 2013). It reports on unconfirmed information about the suspects, from non-official sources. The ECM passive here allows for the journalists to report this story but at the same time capture the uncertainty and their distance to the sources.

A different type of functional account of the passive ECM is presented by Mair (1990) and Noël (1998), who both focus on the discourse function of the passive variant of the ECM as improving “textual cohesion” (Mair 1990: 178) or “thematic progression” (Noël 1998: 1057). Their explanations differ somewhat but what they both argue and show is that the passive allows the most given or topical element to occur in the clause-initial position, which is the preferred position for these elements to occur in. Consider Noël’s example of an ECM passive, followed by a construed active version.

- (138) a. Pub chain trains for Europe
Taylor Walker is benefiting from a customer care programme that it believes will enable it to compete with its European counterparts. Taylor Walker managing director David Longbottom said: "Unlike our European counterparts, in the UK we do not perceive bar work as having professional status. **Bar work is thought to be** transitory, undemanding and unskilled. We are trying to correct this with training programmes and staff recognition schemes" (A7F 695).
b. In the UK **we think bar work to be** transitory, undemanding and unskilled.
(Noël 1998: 1057, his 23a-b)

In (138)a, the passive allows for the most topical element to occur in the subject position of the main clause, providing a fitting transition from the previous sentence. Mair, moreover, provides an explanation for the infrequent occurrence of the active ECM by pointing out not only that the active matrix not only does not provide the benefit for textual fit, but also that the active matrix can be followed by a *that*-clause, a thematically similar but less marked structure (1990: 181). It is the

observations on textual fit by Mair that Los (2005, 2009) uses as the basis for her proposal that the introduction of the ECM passive was caused by the ability of the ECM to create a given, and thus unmarked, subject. We will turn to this topic in Section 3.4.2; the following section first provides a syntactic analysis of the active and passive ECM constructions.

3.4.1.2 A syntactic analysis of ECM

Exceptional Case Marking constructions have always been considered in close connection with raising and control constructions, which all contain a finite main clause and an infinitival subordinate clause. Davies & Dubinsky (2004) provide an excellent overview of how the construction has been analysed in different versions of syntactic theory (focusing on generative theory, but including other frameworks as well). The relevant constructions are given in (139).

- (139) a. Barnett believed the doctor to have examined Tilman.
 b. Barnett seemed to understand the formula.
 c. Barnett tried to understand the formula.
 d. Barnett persuaded the doctor to examine Tilman.

(Davies & Dubinsky 2004: 3, their 1-4)

Example (139)a is the relevant ECM construction, known as Raising to Object preceding the Government and Binding framework, where the crucial idea was that the subject of the lower clause moved (i.e. was raised) to the object position of the higher clause (Postal 1974). Example (139)b shows a so-called Raising to Subject sentence, where the subject of the lower clause moves to the subject position of the higher clause. In both these cases the subject of the lower clause moves to a position in the higher clause. In the last two examples, no movement takes place, but they instead involve control configurations. (139)c is Subject Control, where the subject of the higher clause controls the PRO subject of the lower clause. (139)d is an example of object control, where the object of the higher clause controls the PRO subject of the lower clause.

The analysis of Raising to Object changed in the Government and Binding Framework, and it was in this framework that the term Exceptional Case Marking was first used (see Chomsky 1981: 98). The lower-clause subject is no longer said to move, or be raised, to the object position of the higher clause but now receives case in its original position in the subordinate clause—made possible by the deletion of the clause boundary, otherwise a barrier for movement (1981: 66-68). Because this is a non-standard procedure of case assignment (the standard ways being verb-to-

subject and verb-to-object), and crucially one taking place across a clause boundary, it was called ‘exceptional’.⁴³

The exceptional nature of case assignment in the ECM construction has become less exceptional in Minimalist configurations. Case assignment is replaced by case checking (cf. Section 1.3), which takes place in all instances under specifier-head agreement and triggers movement to the specifier of the relevant projection, with the possibility of covert movement, i.e. movement which is not phonetically realized. Chomsky (1991) proposed two separate projections for the subject (AgrSP) and the object (AgrOP) to check their Case features. In this scenario, the object NP moves to the specifier of AgrOP to check its case features. The consequence of this is that the configuration for case-marking/checking in active ECM constructions is no longer exceptional, since in both an object control and ECM structure, the object now moves to the specifier of AgrOP to check case under specifier-head agreement, and no case assignment takes place across clause boundaries. AgrOP is now not standardly assumed in derivations, but I know of no specific treatment of ECM in recent years which provides a different analysis.

In the passive ECM, unlike in the active version, it is clear that movement has taken place, namely to the subject position of the main clause. Such movement of a lower-clause argument to the subject position of the main clause is reminiscent of raising constructions, and indeed, the structural similarities between the ECM passive and Raising-to-Subject have repeatedly been noted in the literature (e.g. Postal 1974, Chomsky 1981, Warner 1982). As Los (2005: 237) explains, they both show a “mismatch” of case and θ -role assignment: case is received by the argument in a different position (main clause) than the θ -role (subordinate clause).

(140) John is believed to be a liar.

(141) John seems to be a liar.

In both (140) and (141), *John* is thematically an argument of the lower clause *to be a liar*, but *John* receives its (nominative) case in the main clause. The same movement would take place within a Probe-Goal-Agree system with case-checking: the subject position, Spec,TP, needs to be filled and looks for an appropriate element. The only available element is the lower clause argument NP *John*, and *John* subsequently moves to the subject position. Figure 3.8 below illustrates the derivation of the ECM passive, Figure 3.9 shows the derivation of the raising construction.

⁴³ Chomsky (1981: 98) discusses another possible analysis for these sentences, which is similar to Raising to Object—with *Bill* as the object of the main clause governing a PRO subject in the infinitive clause—but he notes that such case marking would still be ‘exceptional’.

Figure 3.8 Analysis of the ECM passive

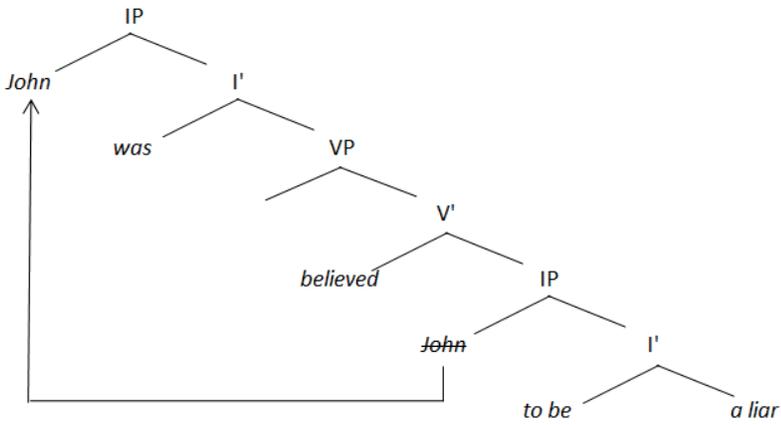
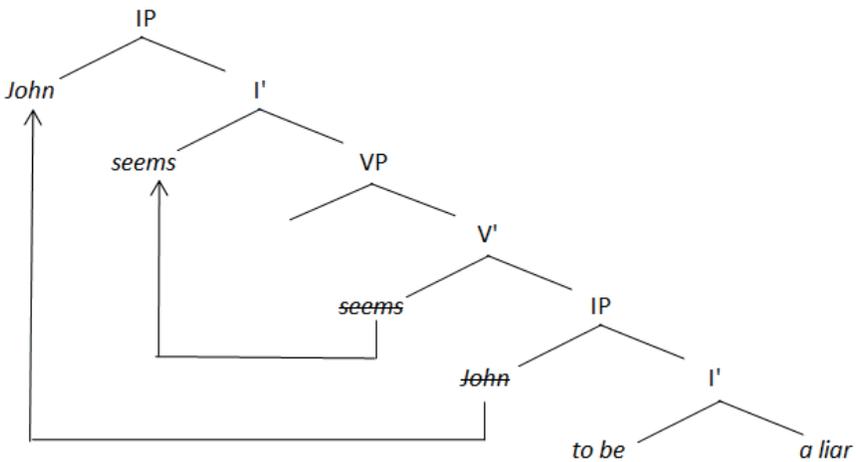


Figure 3.5 Analysis of the raising construction



If we compare the ECM passive to the passive of its close counterpart in the active, the object control construction, it turns out that this derivation is very similar as well.

(142) They commanded John to leave the room.

(143) John was commanded to leave the room.

As in the ECM passive and the raising sentence, there is no external argument in the main clause and the closest available matching argument for the probing Spec,TP position is *John*, the object of the matrix clause. In the resulting structure, the only difference with the ECM passive lies in the original position of what becomes the grammatical subject: this NP either starts out as the subject of the lower clause (in subject raising, and in a raising analysis of ECM) or as the object of the higher clause (in object control, and in a control analysis of ECM). The syntactic operation is the same in both cases: whatever the original position of the NP, this NP is the closest available argument with matching features, which means that this NP becomes the element to fill the Spec,TP position. In other words, the syntactic differences between object control and active ECM, i.e. the difference in original position and the difference in case assignment/case checking, are no longer identifiable on the surface in the passive of these constructions, meaning that in the passive, OC and ECM are even more similar than in the active.

3.4.2 The introduction of the ECM in the literature

Having identified the structural aspects of the ECM and its status in PDE, let us now turn to the introduction of the ECM passive and the questions of when and why it was introduced. In trying to account for the introduction of the ECM passive, all authors mention the influence of the Latin Acl, following Bock (1931) and Warner (1982), but Latin influence is not considered by all to be the only source for ECM. Most importantly, Fischer (1989, 1994) and Los (2005) propose that there is a language-internal development towards ECM which is independent of Latin influence.

3.4.2.1 *Believe* ECM as a loan from Latin

An important early publication on ME ECM is Warner (1982), who provides a detailed study of early ECM occurrences in an attempt to trace and provide an explanation for the introduction of the construction. He discusses non-finite clauses after ‘verbs of knowing, thinking and declaring’ in late fourteenth century sources: the Wycliffe Sermons, the Wycliffe Bible, a collection of examples from Chaucer⁴⁴ and, finally, Visser (1973, §2079, §2081). He distinguishes between two types of sentences: the first is the active ECM (144), abbreviated as *NP to VP*; the other type shows ECM in a context where movement has taken place, including passive ECM as in (145) and *wh*-movement, abbreviated as *NP...to VP*.

- (144) *And þus seiþ Crist, ‘þat he seiþ his apostlis to be his frendis...’*
‘and thus said Christ, ‘that he says his apostles to be his friends’
c1400 (Serm i.170.25; Warner 1982: 136, his 7)

⁴⁴ The collection is based on material from Einkenel (1887), Kenyon (1909), Bock (1931) and Tatlock & Kennedy (1927).

- (145) *þat man is seid to have a fend whom þe fend disseyvþ, as he is seid to have an heed þat is heded bi þis hede;...*
 ‘...as the man who is dominated by a head is said to have this head.’
 c1400 (Serm i.125.2; Warner 1982: 136, his 9)

The first type, showing surface order *NP to VP*, occurs with *boste*, *?confesse*⁴⁵, *coveite*, *gesse*, *graunte*, *?juge*, *seie*, and *wite* in the Sermons. The second type, as in (145), has the same underlying ECM structure but occurs in a different surface order, e.g. by passivization or fronting in relatives, and it is found with *hope*, *seie*, *?unknowe*, *bileve*, *doute*, *suppose*, *teche*, *?telle*, *þenke*, *?trowe* and *?wite* (1982: 135-136). Warner notes, comparing these data to the OED, MED, Visser III.2 and Zeitlin (1908), that the examples in the Sermons for *seie* are “among the earliest ME examples known” (137) and for the other verbs are in fact the earliest instances.

In accounting for the introduction of the ECM, Warner, following Bock (1931), attributes a central role to Latin influence, but with the added observation that the *NP...to VP* type, which includes passive occurrences, seems to be “less restricted in occurrence outside Latin translation[s]” (1982: 136). Warner notes that “almost all ME instances before 1450 are translation[s] from Latin” in Visser’s collection, while 16 out of 23 examples with verbs of knowing and thinking and 5 out of 7 with verbs of declaring in the collection of examples from Chaucer come from his translation of *Boethius* (1982: 141). Additional evidence for the relevance of Latin comes from the Wycliffe translation of the Bible. This translation comes in two versions, as Warner explains, one being a literal translation from the Latin Vulgate and the other a revision: “the evidence presented by Fristedt (1953, 1969) implies that we should view the versions of [the Wycliffe Bible] as stages in a process of revision which (generally) moves steadily away from English styles closely modelled on Latin” (139). Warner finds that translations of Latin *AcI* are often literally translated as ECM in the early version, but are translated with a finite clause in the late version (120 infinitives in the Early Version, 28 in the Late Version; against 8 finite clauses in the Early Version and 99 in the Late Version) (1982: 141). This, he suggests, points to the conclusion that the ECM was not yet completely acceptable in English at that stage and its use in the early stages, including the Early Version, is to be mainly attributed to the Latin influence.

Another important conclusion by Warner is that avoidance of *NP to VP*, i.e. the active ECM without movement, played an important role in the early stages, and, crucially, that it still does so today. This avoidance can be achieved by movement of the NP through, for instance, relativization, topicalization or extraposition. Warner uses the dominance of these ‘avoidance’ structures to explain the introduction of the ECM via a back door, as it were, referring to Bickerton (1973, 1975) for the idea of minimal alterations: “It seems that the Latin accusative and infinitive with such verbs provides English with a ‘target’ which is normally unacceptable and must therefore be approached obliquely by a series of changes which are ‘minimal’ or

⁴⁵ The question marks presumably indicate ambiguous examples in Warner (1982).

those ‘least noticeable’” (157). A sentence with an active ECM with the *NP to VP* is too obviously ungrammatical in these early stages, and the sentences where this order is avoided obscure this configuration.

Although subsequent researchers recognize the relevance of Latin influence, not all of them attribute as much importance to it. Fischer (1989: 160) dismisses Latin as the single cause for the introduction of ECM, based on the observation that other Germanic languages like Dutch and German allow the *Acl* at one stage under Latin influence, but never developed it in the same way as English did. This leads her to conclude that there must be additional factors specific to English that have allowed the ECM to become accepted. Los (2005) also looks for other factors, beyond Latin, but argues more specifically against Warner’s minimal alteration account, stating that it cannot account for the unbalanced distribution of passive and active, and specifically that the passive cannot be considered as a minimal alternation because of the complexity of the movement involved. Most crucially, the minimal alterations only promoted the active ECM in a limited way, because active ECM “never really became acceptable” (2005: 255).

Los (2005: 256-259) proposes, specifically for the passive ECM with *believe*-verbs, that the reason the ECM gained ground in English is the loss of *V2*, following Mair (1990), who observed that their function is to create unmarked themes. The crucial property of ECMs in ME, according to Los, is to create subjects which were previously impossible in the language and which moreover could move given information to the unmarked position for topics after the loss of *V2*, the subject position. The problem with this information-structural motivation is one of timing: we will see in Chapter 5 that the information-structural consequences of the loss of *V2* only start to play a role much later than assumed by Los. This means that the relevance of the ECM subject as an unmarked theme cannot as yet have played as large a role as to lead to the ECM from the earliest instances onwards, although we will see in Chapter 6 that it may have influenced the use of the passive in general in later centuries.

3.4.2.2 Reanalysis of three-place verbs

Fischer (1989) and Los (2005) aim to find a language-internal reason for the introduction and crucially the acceptance of ECM in English, having dismissed Latin influence as the only explanation. Fischer proposes that there is an additional source for ECM in reanalysis of the originally three-place *persuade* verbs as two-place verbs. Her explanation is based on the change in word order from *SOV* to *SVO*, which she locates in “the course of the ME period” (1989: 205), following Canale (1978), Bean (1983) and Koopman (1985). An NP occurring immediately preceding a transitive verb in a finite or non-finite clause in *SOV* OE, she claims, would be analysed as its object, while the same NP in an *SVO* language will be analysed as a subject. Note that this necessarily means an analysis of non-*V2* contexts because *V2* potentially changes the order of object and verb. Fischer illustrates these two possible analyses with the following two examples.

- (146) ... *one het on his gesihðe þone diacon unscrydan*
 ... and commanded in his sight the deacon undress...
 ‘and commanded (someone) to undress the deacon in his presence’
 (ÆCHom I,29 424.11; Fischer 1989: 206, her 83)
- (147) *Moyses forbæd... þa nytenu to etanne*
 Moses forbade the animals to eat
 ‘Moses forbade to eat the animals’
 (ÆLS (Maccabees) 40; Fischer 1989: 206, her 84)

Fischer argues that PDE speakers would analyse (146) in such a way that *þone diacon* is the object of *het*—‘he is commanded to undress’—while OE speakers would analyse *þone diacon* as the object of the non-finite verb *unscrydan*, in which case the deacon is undressed by an unnamed party. Example (147) presents a similar case: PDE speakers will analyse *þa nytenu* as the object of the verb preceding it (the matrix verb *forbæd*), in which case the animals are not allowed to eat, while OE speakers would analyse it as the object of the verb that follows it, the lower verb *etanne*, in which case it is not allowed for an unnamed agent to eat the animals. Fischer points out that the context often makes the interpretation clear. Another aspect that is discussed by Fischer (176-177), with OE examples (with bare infinitives), is that the animacy of the NP can help to distinguish between persuade (three-place) and expect (two-place) verbs: while *persuade* (three-place) verbs invariably occur with animate NPs, *expect* (two-place) verbs allow inanimate NPs.

Los (2005) adopts Fischer’s proposal for a language-internal development towards ECM, but claims that not all three-place verbs (Fischer’s *persuade* verbs) are affected by this development. She distinguishes between so-called verbs of persuading and urging, which are always three-place verbs, and verbs of commanding and permitting, which can be both three-place and two-place (2005: 237). She proposes that the reanalysis of commanding and permitting, originally three-place but innovatively two-place with a *to*-infinitive, dates back to the M1 and M2 periods of the PPCME (1150-1350), based on an analysis of 8 verbs of causing, commanding and permitting: *biddan* ‘ask’, *commaunden* ‘command’, *don* ‘do’ (causative), *haten* ‘order’, *leten* ‘let’, *maken* ‘make’, *ðolien* ‘suffer’, *suffren* ‘suffer’. She analyses 13 of the total number of examples with these verbs as ECM rather than object control structures, as presented in Table 3.5 below.

Table 3.5 Object Control *NP to VP* complements versus *to*-infinitival ECM after verbs of causing, commanding and permitting (Los 2005: 243, her table 9.3).

	M1	M2
<i>NP to VP (control)</i>	75 (93%)	38 (84%)
<i>NP to VP (ECM)</i>	6 (7%)	7 (16%)

Los’s most important test for analysing a sentence as ECM is the animacy of the NP in the *NP to VP* sequences, following Postal (1974) and Fischer (1989). What she

finds is that the NPs in the *NP to VPs* with verbs of commanding and permitting are increasingly inanimate, leading to the conclusion that they are no longer recipients and therefore no longer arguments of the higher verb, but only of the lower verb (2005: 239-241). This in turn leads to an ECM rather than an object control reading.

Los gives two examples of these commanding and permitting verbs reanalysed as ECMs, given in (148) and (149) below.

(148) *and **makeþ** to comen al out of smak. al þet me wes ywoned byuore to louie*
 ‘and he makes completely insipid everything that one used to love’
 c1340 (Ayenbi, I, 106.146; Los 2005: 242, her 28)

(149) *þou **comaunded** by comaundement to be greteliche kept*
 ‘you commanded your commandment to be carried out to the letter’
 (Earlps, 145.2145; Los 2005: 242, her 28)

These examples are clearly ECM constructions: in both cases the relevant NP is inanimate and so the sentence can never be interpreted as an object control structure. At the same time, it is interesting to note that these examples occur in similar contexts to the first examples of the *believe*-ECM: the first example avoids surface *NP to VP*, because the ECM object occurs after the infinitive, and both examples occur in translated texts, the *Azenbite* from French, *The Earliest Complete English Prose Psalter* from Latin and French. The example from the *Early Psalter* is interesting in other ways as well: it seems to be a quite literal translation from the Latin *Vulgate*, as (150) shows.⁴⁶

(150) *Tu mandasti mandata tua custodire/custodire nimis.*⁴⁷
 You have-commanded commandments your to-keep/to-be-kept diligently.
 (Vulgate, Psalms 118:4)

The use of the passive infinitive leads to an ECM interpretation of the example and avoids an ambiguity in the status of the NP as either the object of *command* or *keep*. Because *your commandments* is inanimate, it cannot thematically be the object of *command*, which would be a possible reading for a speaker of an OV language in the active version *You have commanded your commandments to keep diligently*, similar to example (147) above. Interestingly, the *Wycliffe Bible*, another translation from the Latin *Vulgate* but two centuries later than the *Early Psalter*, has a less literal translation, which is in addition unambiguous with respect to the status of the NP, given below.

⁴⁶ The Vulgate is the source for the Early Psalter (both manuscripts), although there may be differences with the modern versions of the vulgate (Bülbring 1891: introduction, p. x)

⁴⁷ There is a difference between versions of the Vulgate with respect to the choice of an active or passive infinitive.

(151) Thou hast commanded; that thy behests be kept greatly.

(Wycliffe Bible, Psalms 119:4)

The Latin (passive) infinitive is here translated with a *that*-clause with a subjunctive passive infinitive, meaning that it is not possible to interpret this as an ECM example. The translation in the English *King James Version* (1611)—which uses the Hebrew source text rather than the Latin *Vulgate*—has an unambiguous object control sentence, with the NP of the *NP to VP* as a clear recipient (i.e. argument) of the higher-clause verb.

(152) Thou hast commanded us to keep thy precepts diligently.

(King James Version, Psalms 119:4)

Unfortunately, there are no occurrences of *command* in the PPCME2 with an active infinitive that could shed more light on the behaviour of *command*, probably due to the fact that the PPCME2 is small for purposes of lexical searches and *command* is an early ME loan.

Most importantly, these examples show that ECM-like sentences are already attested earlier than the examples with *believe*. Los therefore pushes the introduction of the type of ECM after verbs of commanding and permitting forward, i.e. before the introduction of the *believe*-ECM. Fischer (1989: 205) already located the change following the OV to VO change, “in the course of the ME period”, and Los’s examples above date from the twelfth and thirteenth century. This supports the analysis by both Fischer and Los that the *believe*-ECM and *want*-ECM are two separate developments, and that their introductions should be taken to be two separate developments as well. While the timing is difficult to prove because of the scarcity of examples in M1 and M2 and the ambiguity that is inherent in the construction, the two developments may have reinforced one another once both constructions were attested, as Los points out (1999: 293). It is especially interesting in this regard that similar factors avoiding surface *NP to VP* are attested in the early examples of both types of ECM. One of the questions that remain is whether the passive ECM examples show a similar development, originating from two different sources of ECM.

3.4.2.3 A separate passive loan (Noël (2008))

Noël (2008) presents an account which disconnects the passive ECM completely from the active and which can potentially explain the unbalanced distribution that has been found from the introduction onwards. He claims that the ECM passive with verbs of thinking and declaring is not in fact a passive of the active ECM but that it is a separate loan from Latin.⁴⁸ Latin, according to Noël, had both an *Acc* (*accusativus*

⁴⁸ Note that this seems to be a crucially different account than his grammaticalization account presented in earlier publications (1998, 2001).

cum infinitivo) and a so-called Ncl (nominativus cum infinitivo). This Ncl is a passive with an infinitive and has what Noël calls an evidential reading, which he describes as having “an (epistemic modal) auxiliary-like function” (2008: 316), cf. Section 3.4.1.1. He provides the following fifteenth-century examples from the OED where a Latin Ncl is translated with *is said to* in English.

- (153) *Puluer of vitriol combuste . . . is seid to availe agayns polipe in þe nose.*
powder of combusted vitriol [or ‘burnt sulphuric acid’] ... is said to help
against polyps in the nose
(J. Arderne, *Treatise fistula*, Sloane 6:79)
- (154) *Margaryte, whyche gemme is white lytyl and vertuose . . . The vertu of thys
Stone is sayd to be ayenst effusyon of blood.*
Margarite, which gem is white, little and virtuous ... the virtue of this stone is
said to be against effusion of blood
(Caxton, *The golden legend*, 214/2 A)
(Noël 2008: 324, his 24-25, translations mine)

Noël’s scenario in which the Ncl is a loan from Latin is only based on a handful of examples and unfortunately, he seems to have overlooked Warner’s (1982) findings that the passive ECM is also found outside Latin translations from the first instances onwards, in contrast to the active ECM. There is an additional complication in his story—which is also problematic to a certain extent for the grammaticalization account—in that Noël claims that the surface *be-passive participle-to* form can have three different readings, which are however not always easy to distinguish. He illustrates his three categories with the following examples.

- (155) Sometimes authorities are understood to be limited by the kinds of acts which they can or cannot regulate (given some restrictive ways of classifying acts). In this book authorities are said to be limited also by the kinds of reasons on which they may or may not rely in making decisions and issuing directives, and by the kind of reasons their decisions can pre-empt. (ANH 148)
- (156) Horatian metres both dictate and accommodate Horatian syntax; in every poem there are striking effects of word order which, on the one hand, **may be said to have been contrived**, or willed by the poet, but on the other can be seen as arising from metrical necessity. (J7P 44)
- (157) INDEPENDENTS GET BREAK FROM BRAVO
by David Goymour
BRAVO, a new on-screen booking system which puts British hotels and tourist attractions on travel agents’ counters, has entered its launch phase. It is said to offer independent hotels the kind of exposure which hotels in big groups

can derive from international booking systems—owned, typically, by the big airlines. (AOC 456)

(Noël 2008: 317, his 9-11)

According to Noël, example (155) is a regular passive: the verb *say* refers to an actual speech act, namely in the book. Examples (156) and (157) are different in that they do not refer to an actual speech act. Example (156) shows an example of what Noël calls the ‘descriptive Ncl’ because it shows “one of the possible descriptions” (318) of the “striking effects”. Example (157) shows what Noël calls the “evidential Ncl” because the writer indicates that he has a source for the information, but crucially “sheds responsibility for the truthfulness of this information” (318). If this distinction is hard to make in PDE, it is likely to be even harder for ME examples. Despite the complication of the difficult distinctions between the categories, I believe that the possible evidential function of the ECM passive in PDE is clear. Whether this is due to a grammaticalization or a loan remains to be seen, but the evidence for a separate loan seems far from convincing.

3.4.2.3 Conclusion

This discussion of proposals in the literature for the introduction of the ECM construction, both active and passive, has shown that there are two different types of ECM that were introduced in ME: one is the ‘genuine’ ECM with *believe*-verbs, i.e. two-place verbs; the other is ‘reanalysed’ ECM with *want*-verbs, i.e. verbs that are originally three-place, but occasionally occur as two-place verbs under an ECM configuration. This also means that we need two different explanations for their introduction. For the *believe*-verbs, Latin influence, as shown by Warner (1982), is recognized as an important factor, but it is not considered to be the only relevant factor by all authors, with Fischer (1989) and Los (2005) aiming to find language-internal reasons. For Fischer, this is the related reanalysis of object control to ECM which she ascribes to the word order change from OV to VO; for Los it is the loss of V2. Warner and Fischer mostly consider active and passive as one construction in their discussion, while Los (2005) proposes the V2 factor specifically for the ECM passive, in an attempt to explain why the ECM passive was accepted in English more readily than the active. Noël (2008) also presents a specific proposal for the passive ECM, but there are problems with his data. Table 3.6 below summarizes the different types of active ECMs and their introduction according to the literature.

Table 3.6 ECM and related constructions in the active

Verbs	Type	Time of introduction
Perception verbs	Acl	Since OE with bare-infinitive
Persuading and urging	Object control	Since OE with <i>to</i> -infinitive

Commanding and permitting	Object control	Since OE with bare-infinitive and <i>to</i> -infinitive
Commanding and permitting	ECM	Los (2005): from early ME onwards, only with <i>to</i> -infinitive
<i>Believe</i> -verbs	ECM	Late ME with <i>to</i> -infinitive, first in translations from Latin

3.4.3 The ECM passive in the PPCME2

The previous section has underlined the importance of distinguishing between groups of verbs, and the distinction between the different types of ECM that are introduced in the ME period. The aim of the present section is to investigate whether the PPCME2 provides new evidence for the scenarios presented in the literature and whether these data can shed new light on the introduction of the passive ECM specifically, in relation to the introduction of the active ECM. In this section, I present three observations from the study in the PPCME2 of the passive ECM and related passives: (i) *believe*-ECM is introduced with both an active and a passive form from the first examples onwards, but the passive dominates; (ii) with the verbs of commanding and permitting, the introduction of the passive is the result of the introduction of the recipient passive, and these passives provide only little further evidence for the reanalysis scenario; and (iii) there are several passives which are structurally alike to the ECM passive, and which, due to their structural similarity as well as their frequency, I argue play an important role in the development of the ECM passive.

3.4.3.1 The ECM active and passive with *believe* verbs

Unfortunately, the ECM passive is not annotated separately in the PPCME2, nor does it have a specific structure on the basis of which it can be selected.⁴⁹ The following query provides only an initial selection: it selects all passive clauses (i.e. a clause with a form of *to be* and a passive participle), which should in addition dominate an infinitival subordinate clause.

<i>Query 3.4 ECM passives</i>
a. IP-MAT* IP-SUB* idoms *BEP* BED* BEN BAG BE
b. IP-MAT* IP-SUB* idoms *VAN*
c. IP-MAT* IP-SUB* idoms IP-INF*
d. *BEP* BED* BEN BAG BE precedes *VAN* precedes IP-INF*

The last line (3.5d) stipulates that the order of elements should be as follows: *to be*, past participle, infinitival subclause. The results of this query include not only ECM passives, but also passives of object control; the different types of passives were

⁴⁹ It is true that some distinctions are in fact made in the corpora—such as whether or not the infinitive has an (empty) subject node—but the annotation is not entirely systematic.

selected by hand. I also carried out a number of lexical searches based on the list of verbs that occur with the ECM from Warner (1982), in order to make sure that I did not overlook any examples that were annotated differently.

The ECM passives are not frequent in the PPCME2, with only 9 examples, and they do not occur before the M3 period (1350-1420), corresponding to the previous findings in the literature. The first three examples in the corpus, given in (158)-(160), are from *Purvey*, which is dated at c1388 and crucially not a text that is translated from Latin.

(158) *and in the xxxiiij. c. of Ezechiel, he is seid to feede him self and not the scheep; and in the 34th chapter of Ezekias, he is said to feed himself and not the sheep*
c1388 (cmpurvey,31.1518)

(159) *and these ben seid to be maad of nougt in to prelatis, and these are said to be made from nothing into prelates⁵⁰*
c1388 (cmpurvey,I,31.1548)

(160) *also he is seid to seke his owne profitis temporal, also he is said to seek his own [temporal benefits]*
c1388 (cmpurvey,I,31.1519)

An interesting observation about these examples is that (158) does not have the ECM evidential reading but is a literal passive of ‘someone specific said x’ rather than the modern type ‘it is allegedly said somewhere by someone but I do not know who or where and it may not be true’. This is also a likely reading for (159). This would mean that the *believe*-ECM really starts out as a passive of an active model, rather than a separate construction, as Noël (2008) proposes. The fact that these are examples with *say* and no translations from Latin also means that Noël’s proposal of ECM with *say* being a separate loan from Latin cannot hold.

The following table shows the dates, texts, number of occurrences and the verbs of all ECM passives found in the corpus.

Table 3.7 *Believe*-verbs with ECM in the PPCME2

c1388	CMPURVEY	3	<i>said</i>
c1400	CMWYCSEY	2	<i>said</i>
?a1425 (c1380)	CMBOETH	2	<i>deemed</i>
a1450 (c1391)	CMASTRO	2	<i>imagined</i>

Note that of the four texts, only the *Boethius* text is a translation from Latin, confirming Warner’s (1982) observation that the passive is less restricted to Latin translations. Two of the examples in fact are from the same corpus as Warner

⁵⁰ A prelate is “an ecclesiastic of high rank” (MED)

(1982), the Wycliffe sermons. The same verbs occur more than once in the texts, and all five earliest examples (ordered on the basis of manuscript date rather than date of composition) have *say* as the main verb. Examples (161) and (162) illustrate the two other verbs that are found in the corpus.

(161) *For thilke same blisfulnesse is demed to ben sovereyn suffisaunce,*
for this same blissfulness is deemed to be sovereign sufficiency
?a1425 (c1380) (cmboeth,433.C2.194)

(162) *Thys forseide grete pyn in manere of an extre is ymagyned to be the Pool Artik in thyn Astralabie.*
this aforementioned great pin in manner of an axletree is imagined to be the arctic pole in your astrolabe
a1450 (c1391) (cmastro,665.C2.93)

It is not clear from the context of either example whether these represent instances of an evidential reading or whether they refer to an actual speech act.

Most of these verbs do not occur in the active ECM in this corpus. Query 3.5 below was used to search for these active occurrences.

Query 3.5 Active ECM
a. IP-MAT* IP-SUB* idoms *VBD* *VBP* VB
b. IP-MAT* IP-SUB* idoms IP-INF*
c. IP-INF* idoms TO
d. IP-INF* idoms !FOR
e. VBD* VBP* iprecedes IP-INF*

As for the passive ECM, the coding in the PPCME2 is not consistent, and so, like Query 3.4, this query provides only an initial selection of possible relevant sentences and the relevant examples were selected by hand. Again, I added lexical queries to search for specific verbs in the active. The results of these queries show that *imagine* does not occur as an ECM active in the PPCME2. *Seie* occurs in the active ECM in Warner's (1982) examples (from the Sermons), but not in the PPCME2, and *demen* does not occur in the active ECM in the PPCME2 either. This confirms the dominance of passives over actives that has been noted before, from the earliest examples onwards. At the same time, it shows that some verbs that do not allow the active in PDE did allow it in earlier stages, which means that the distribution of active and passive may have changed throughout the centuries.

3.4.3.2 Verbs of commanding and permitting

With regard to the verbs of commanding and permitting—the group that Los (2005) proposes show the possibility of reanalysis of object control to ECM—there are two questions: (i) Los (2005) shows that these verbs take a dative recipient and so there

is the same question as for the recipient passive: when do recipient subjects of passives become available? And (ii), do the passives of the verbs of commanding and permitting show any sign of the proposed reanalysis that has taken place in the active?

Los (2005) proposes that the original patterns of the verbs of both persuading and urging and commanding and permitting “persist through history”: the former take accusative objects, while the latter take dative objects (2005: 249). Because they take accusative objects, the verbs of persuading and urging can already passivize in OE. The fact that the verbs of commanding and permitting also start to occur in the passive towards the end of the ME period can be explained with reference to the introduction of the recipient passive with two nominal complements, as described in Section 3.3. Los finds little evidence of these recipient passives with one nominal complement and one clausal complement in the early ME periods (PPCME2, M1 and M2), confirming Allen’s (1995) earlier observations. As an exception, Allen points to *teach* and *thank* (1995: 401), but she concludes that other examples show that certain verbs could follow two patterns and assign different cases in clausal than in regular ditransitives (i.e. with two NPs). In the absence of convincing, i.e. unambiguous, early examples of recipient passives with verbs of commanding and permitting, it seems safe to conclude that the introduction of recipient clausal ditransitives is indeed connected to the introduction of recipient passives with two nominal complements in the late fourteenth century. The introduction of the passive with verbs of commanding and permitting, then, is not connected to the introduction of the *believe*-ECM passive.

With respect to the question of reanalysis of object control to ECM of the verbs of commanding and permitting, it is somewhat difficult to establish whether reanalysis has taken place in passives that occur with the verbs of commanding and permitting because of the syntactic similarities between ECM passives and passives of object control. As we saw in Section 3.4.1.2, some of the crucial differences between object control and ECM seem to be obscured in the passive versions of these constructions. One of the clues that is used by Los to determine the status of the object/subject NP in the active is animacy: if the NP in an *NP-to-VP* sequence is inanimate, an object control is not possible, while if it is animate, both OC and ECM are possible. This difference persists in the passive, as illustrated in (163).

- (163) a. You commanded [your commandment to be carried out to the letter].
 b. *You commanded [your commandment] [PRO to be carried out to the letter].
 c. [Your commandment_i] was commanded [[t_i] to be carried out to the letter].

The ECM analysis in (163)a is the only possible analysis of this sentence, with the object control interpretation in (163)b being impossible with an inanimate NP. When this sentence is passivized, the result is of course a passive with an inanimate subject, and the structure is as in (163)c.

The passives of the surface structure *to be-passive participle-to* were selected with the same query that was used to select ECM passives (Query 3.5), and they were categorized according to the type of verb in the matrix clause. In total, I found 17 examples with verbs of commanding and permitting in the PPCME2, based on Los's lists of these verbs (2005: 102). The results of the analysis are given in Table 3.8.

Table 3.8 Animacy of subject of passive with verbs of commanding and permitting

Verb	Animate	Inanimate
<i>command</i>	7	-
<i>labour & command</i> (conjoined)	1	-
<i>ordain</i>	2	1
<i>bind</i>	1	-
<i>assign</i>	1	-
<i>charge</i>	1	-
<i>make</i>	1	-
<i>constrain</i>	1	-
<i>pardon</i>	1	-

Table 3.8 shows that virtually all subjects are animate, providing us with no further evidence that these examples are ECM structures. There is one example with an inanimate subject, with ECM as the only possible analysis, given in (164).

- (164) *And by the kyng and hys counselle a Parlyment was ordaynyde to be-gyn on Syn Leonarde ys day nexte folowyng.*
 and by the king and his council a parliament was ordained to begin on St. Leonard's day immediately following
 c1475 (cmgregor,195.1525)

Because of the inanimate subject, the active version cannot be an object control (cf. *the king ordained a Parliament to begin*). While it does not provide additional evidence for reanalysis, the dominance of animate subjects among the passives is not unexpected and can be explained with reference to other factors: while objects have no restrictions on animacy, subjects are more generally animate, with the result that subjects are more likely to be animate NPs, both of active and passive clauses. This could explain the small number of examples with inanimate subjects.

Another indication for an ECM or OC reading is the type of infinitive: a passive infinitive makes an ECM analysis the only possibility, as we have seen in example (149) from Los. It has often been observed that a considerable number of potential ECM examples occur with such a passive infinitive (Fischer 1989: 207-208, Los 2005, Warner 1982), as in (164).

- (165) & [the pope] *commanded* pat psalm to ben seyd every day at pryme
 and the pope commanded that psalm to be said every day at Prime [first
 hour of daylight]
 ?a1425(c1400) (cmmandev,96.2331)

When the infinitive is passive, the NP immediately following the matrix finite verb cannot be a recipient of that verb: *pat psalm* is not the recipient of the main clause verb (it is not the psalm that is ‘commanded’ to do something), which means that this cannot be an object control structure and the ECM reading is the only reading that is available. This means that a passive infinitive can be another clue for reanalysis, following both active and passive main clauses, but passive matrix clauses with passive infinitives do not occur in the corpus.

The clauses with passive infinitives were selected with the following queries. Note that no distinction is made between different types of verbs – so they could contain both commanding and permitting verbs well as persuading and urging verbs; although verbs of persuading and urging do not allow ECM, examples with passive infinitives are possible (see below).

Query 3.6 Active object control with a passive infinitive

- a. IP-MAT*|IP-SUB* idoms *SBJ*
- b. IP-MAT*|IP-SUB* idoms *VBD*|*VBP*|VB
- c. IP-MAT*|IP-SUB* idoms IP-INF*
- d. IP-MAT*|IP-SUB* idoms NP-OB*
- e. *VBD*|*VBP*|VB precedes NP-OB*
- f. NP-OB* precedes IP-INF*
- g. IP-INF* idoms *VAN*
- h. IP-INF* idoms !NP-SBJ*
- i. IP-INF* idoms !FOR

Query 3.7 Active ECM with met passive infinitive

- a. IP-MAT*|IP-SUB* idoms *SBJ*
- b. IP-MAT*|IP-SUB* idoms *VBD*|*VBP*|VB
- c. IP-MAT*|IP-SUB* idoms IP-INF*
- d. IP-INF* idoms NP-SBJ*
- e. IP-INF* idoms *VAN*
- f. *VBD*|*VBP*|VB precedes IP-INF*

There is only one relevant example in M1 and M2, which is the example given earlier by Los (2005), see (149). Other possible early examples (on the basis of composition date rather than manuscript date) are given below.

- (166) *And hardely I dar boldely seyn þat he wyl not denye his feet to þe, þat art a mayde, wan he **grauntede hem to be kyst** of a synful womman.*
and hardly I dare boldly say that he will not deny his feet to thee, who are a maid, when he granted them to be kissed by a sinful woman
c1400 (cmaelr3,43.502)
- (167) *Noght forþi þan he **suffers þam to be temped** on sere maners, boþ wakand and slepand.*
nevertheless he then suffers them to be tempted in many ways, both awake and asleep
a1450 (?1348) (cmrollep,90.438)

From M3 onwards, the numbers of relevant examples increase, but this is also the period during which the number of ECM examples increases in general. Unfortunately, it is not possible on the basis of these queries to determine the exact number of active OC and ECM structures and thus to determine the percentage of passive infinitives among them.

3.4.3.3 Related passives

There are several structures available in OE and ME which are both on the surface and on a structural level very similar to ECM passives and may have provided a model for the ECM passive. Despite their differences in the derivation—which we have seen in Section 3.4.1.2 are somewhat obscured in the passive (the discussion of whether an ECM argument starts out as subject of the lower or object of the higher)—the surface structure is similar in all these passives: subject-*be*-passive participle-*to*-infinitive. I would like to propose that the existence of these structures meant that the ECM passive structurally was not an obvious innovation, because the new ECM configuration could go unnoticed, and that this is the reason why it came to be accepted easily in English, in contrast to the marked active ECM structure. The existence of these patterns has already been noted in the literature, including their structural similarity to ECM, but what the study of the PPCME2 especially makes clear is that the frequency of these other types of passives must have played an important role in making the ECM passive a less suspect construction than the active version.

Verbs of persuading and urging

The first group that is relevant because it has a similar surface structure to the ECM passive consists of the verbs of persuading and urging. As Los (2055: 249) explains, these are always unambiguously three-place (i.e. no reanalysis to ECM has taken place) and take accusative objects, so they could passivize already. At the same time, Los remarks that “[t]here are occasional examples in ME in which a passive *to*-infinitive after verbs of persuading and urging creates structures strongly reminiscent of ECM” (249). She gives the following example.

- (168) *thei be thret to be slayn or presoned*
 they are threatened to be slain or presoned
 ‘they are threatened with death or imprisonment’
 (PL, Davis 1971: 194, ll. 7–8; Los 2005: 249, her 45)

Similar to what we have seen in the previous section, the infinitive is passive, which means that the relevant NP cannot be a recipient of the higher verb, which makes it look like an ECM example, although persuading and urging do not normally allow ECM. This again adds to the ambiguity that already exists between these structures in that some of the distinctions in the active may be blurred in the passive.

Because these verbs are three-place and take accusative objects, passivization was already possible in OE and the PPCME2 should contain examples from the earliest ME periods onwards. They were selected with the same query that was used to find ECM passives (Query 3.5), and the relevant sentences were identified on the basis of the list of verbs from Los (2005). Relevant examples are relatively easy to find from the earliest periods onwards; the following examples show a selection of these early examples.

- (169) *Hys hert **is diȝt to hope** in our Lord,*
 his heart is ordered to hope in our Lord
 c1350 (cmearlps,140.6114)
- (170) *asswa se ȝe maȝen iseon. water hwenne me Punt hit & stoppeð hit bifore wel
 þt hit ne maȝe dunewart þenne **is hit ined aȝein forto** climben upwart.*
 also so ye may see water when men dam it, and stop if before well that it not
 can downward then is it forced again to climb upwards.
 ‘Just as you can see water, when someone dams it and stops it up firmly so
 that it cannot go downward, having to go on rising upward’
 c1230 (cmancriw,ll.59.596) (translation from Savage & Watson 1991)
- (171) *þerafter at ȝork þey **were compelled to defende hem self;***
 thereafter at York they were forced to defend themselves
 a1387 (cmpolych,VI,339.2484)
- (172) *and **was counselyd to remove thens to a towne xx myle thens that is callyd**
 Solye.*
 and was advised to remove thence to a town 20 miles thence that is called
 Solye
 c1450 (1438) (cmedmund,173.288)
- (173) *His sones, anon as þei were of age, **were lerned to ride and to exercise hemself**
 in dedis of armes.*
 his sons, as soon as they were of age, were taught to ride and to exercise
 themselves in deeds of arms (a1464 (cmcapchr,83.1549))

- (174) *Thys yere, a-bute Mydsomyr, a the ryalle feste of the Sargantys of the Coyfe, the Mayre of London was desyryde to be at that feste.*
this year, about Midsomer, at the royal festival of the Sergeants of the Coif,
the mayor of London was desired to be at that festival
c1475 (cmgregor,222.2187)

The fact that these passives with persuading and urging, in an object control configuration, were already used in English before the introduction of the ECM passive, together with the syntactic similarities between object control and ECM in the passive, means that the passive ECM can be seen as a minimal alteration.

Fixed adjectival constructions

Another construction which is relevant to the ECM passive because it also has a similar structure (by virtue of being like raising verbs, see e.g. Postal 1974) are what I will call the fixed adjectival constructions. These sentences contain a passive participle which seems to have lexicalized into a fixed construction where the status of the passive participle as verbal is no longer clear (cf. the distinction in Chomsky (1981) between syntactic and lexical passives). Examples (175) and (176) illustrate.

- (175) & *wes iwunet ofte to cumen wið him to his in. & iseon his dohter.*
and was wont often to come with him to his lodgings and see his daughter
c1225(?c1200) (cmjulia,96.12)
- (176) *He was wone to sey, 'I entred euel,*
he was wont to say 'I entered evil,
a1464 (cmcapchr,103.2212)

The reason that they are relevant for the ECM passive is not just that they are alike in surface structure (*be-participle-to*), but that they are also very frequent in the corpus. Examples are present in all subperiods of the PPCME2 and they represent a considerable number of all *be-participle-to* surface orders: in M3, for instance, they make up more than half of the *be-participle-to* examples.

While it is more or less clear that these types of constructions are fixed expressions in PDE, this is less obvious for the ME period, as they might still be in the process of becoming fixed: grammaticalizing from a full productive verb into an adjectival participle with a specific function. However, as we will see, there is enough evidence to suggest that the most common examples, those with *wont*, *bound* and *used*, are indeed already lexicalised as adjectives and do not function as verbal participles anymore.

The most common participle, *wont*, hardly occurs with the meaning of 'having a habit' in the PPCME2. The MED gives an active entry for the verb *wonen*, but these represent examples with a specific meaning of *living somewhere, dwell or residing*. It has a separate entry (4b) for passive uses of *wont*, which already have the same interpretation as in later examples: "to tend by habit (to do sth.), be accustomed;

also, be accustomed (to have sth.)". It also has a separate entry for the adjectival use: "customary, accustomed". Moreover, most of the examples in the PPCME2 fit this interpretation, and indeed, the meaning of *dwelling* or *living* does not occur in the active, nor is there an active equivalent of *be accustomed to*.

The case for *wont* seems quite clear, but for other verbs it is more difficult to say whether they are indeed already grammaticalized in the ME period. The two most common verbs are given in (177) and (178).

(177) *and be vsed to swere horrybull opys by Godys sydys and his blod, and vmbreyden God of his passyon,*
and are used to swear horrible oaths by God's sides and his blood and reproach God for his passion
a1500(a1415) (cmmirk,132.3512)

(178) *bis prayore eury Cristen man is bondon to conne, and to preye to God by bis preyour,*
this prayer every Christian man is bound to know, and to pray to God by this prayer
c1450(c1425) (cmroyal,12.75)

These sentences are fixed adjectival constructions in PDE, similar to *be supposed to* (cf. Noël & Van der Auwera 2009 and Mair 1990), but their status in ME is less clear. For instance, *bound* can at this time also be used as in (179), where it is possible to interpret it as a passive with an implied agent rather than only the adjectival meaning.

(179) *bese vij verkes þou arte bondon to fulfill by verke and dede giff þi powere be,*
...
These 7 works you are bound to fulfil by action and deed if your power is.
'You are bound to fulfil these 7 works by action and deed if it is within your power'
c1450(c1425) (cmroyal,19.167)

Even though *wont* and some of the other participles are grammaticalized and in that sense are not really connected to the passive ECM, which is more clearly a verbal passive rather than an adjectival one, I think they are relevant because they provide a template for the ECM, not in the least because of their frequency. As such, the ECM passive represents only a minimal alteration, and not a full structural innovation, from these fixed adjectival constructions and the object control passives.

3.4.3.4 Conclusion

The data presented in this section focused on two lines of investigation: (i) the reason for the introduction of passive ECM with *believe*; and (ii) the proposed

reanalysis of object control to ECM with verbs of commanding and permitting. The corpus provides little to no evidence for the second point, as the findings confirmed earlier proposals that the passives with the verbs of commanding and permitting—taking dative objects under an object control configuration—became possible with the introduction of the recipient passive. Only limited evidence is available in the passive versions that reanalysis has taken place. With regard to the introduction of the passive ECM with *believe*, the examples from the present study suggest that it started out as a regular passive of the active ECM after that had been introduced, i.e. the active and passive are introduced around the same time. The reason that the passive soon became more frequent in English is, I believe, due to the fact that the passive was not an obvious structural innovation. The passives of persuading and urging, of commanding and permitting, as well as the fixed adjectival constructions that were already available in English are both surface-wise and syntactically very similar to the ECM passive and provide a model for the ECM passive. Where the active ECM met with restrictions—it was limited to Latin-influenced texts, and had the apparently problematic *NP to VP* sequence—the passive did not meet with such restrictions. The availability of similar constructions which could function as a template seems to have paved the way for the ECM passive.

3.4.4 Summary and conclusions

This section has investigated the introduction of the ECM passive in ME. It took into account the introduction of the active as well as the passive, because the construction as a whole was introduced around this time. The PPCME2 results confirm earlier findings that the ECM passive with *believe*-verbs was introduced towards the end of the fourteenth century, while it could not only provide limited further evidence for the proposal in the literature for a reanalysis of verbs of commanding and permitting from object control to ECM. I argued that while the ECM passive was introduced as a regular passive of the active ECM, its less marked structure explains the dominance of passive over active that has been observed from the first occurrences.

The corpus study of the PPCME2 confirmed earlier findings in the literature, most notably Warner (1982), that the ECM with *believe*-verbs, both active and passive, was introduced in the late fourteenth century, although there were only a handful of examples of passives. Not all of these examples were translations from Latin, again confirming Warner's findings that the passive is less restricted in occurrence to Latin contexts. This also meant that Noël's (2008) proposal of the passive ECM as a separate loan from Latin cannot hold. A discussion of Fischer (1989) and Los (2005) led to a separate line of inquiry with the verbs of commanding and permitting. I have only found limited further proof in the PPCME2 for the proposed reanalysis of OC to ECM. However, I would like to argue that this reanalysis plays a minimal role in the development of the passive ECM (both *want* and *believe*) because of the structural similarities between the passives of these constructions (as pointed out earlier in the literature).

These structural similarities are also the reason for the dominance of the passive, not some other reason to do with its introduction, such as a direct loan from Latin, or a separate development: the active examples with *say* show that the active is to some extent possible and it really is a passive of an active, i.e. they were introduced at the same time as one construction. The dominance of the passive is due to the fact that there are few structural differences between the passive ECM and related, but crucially structurally distinct, constructions in the active. This means that a model for ECM was already present and that the passive did not meet the same restrictions as the active ECM. The question that remains is how these developments continue after the first introduction; we will turn to this in Section 6.3.

Similar to the recipient and prepositional passives, the ECM passive was introduced before the loss of V2; most importantly, the information-structural motivation for ECM—as formulated by Mair (1990), Noël (1998) and Los (2005)—could not have played a crucial role at the stage of the loss of V2 that the ECM passive was introduced. In contrast to the recipient passive and prepositional passive, the introduction of the ECM is not directly related to other changes in the language in the same period, but nevertheless builds closely on specific constructions available in the language in the relevant period.

Section 3.5 Conclusion

This chapter has investigated the introduction of the new passives, investigating both the time of their introduction, as well as the reasons for their introduction and the extent to which they represent structural innovations. The main conclusions are that (i) the time of introduction may be even closer than previously assumed, in the second half of the fourteenth century, with the prepositional passive as the earliest, but all clearly before the loss of V2 was under way; (ii) there are small-scale changes in the language that are relevant to the introduction of the prepositional and recipient passive and the ECM passive is really a passive of the newly borrowed active ECM; and (iii) the three introductions are independent developments and cannot be interpreted as a single structural innovation in the passive.

Section 3.2 investigated the introduction of the prepositional passive and I argued that the time of introduction for the prepositional was later than previously proposed, most notably by Van der Gaaf (1929), Visser (1973) and Denison (1985, 1993). The early thirteenth-century examples are problematic, and it is only in the second half of the fourteenth century that less ambiguous examples start to occur, and more of them. I showed that reanalysis of verb and preposition as a unit was an important factor from the first examples onwards: many of the early examples suggest that reanalysis had taken place, for instance because the verb and preposition are a translation from a single Latin verb. This reanalysis in itself cannot explain the introduction, which is why I proposed that the introduction is ultimately due to a series of minimal alterations, leading to increased acceptability: the

examples in the PPCME2 show that many of the early prepositional passives are found in relative clauses, where stranding was already possible, even in OE.

Section 3.3 investigated the introduction of the recipient passive, which I argued was caused by the loss of the dative-fronted passive, in turn related to the bundle of changes that make up the loss of V2. The corpus study of the PPCME2 did not give reason to reconsider earlier conclusions by Allen (1995) that the recipient passive was introduced in the late fourteenth century. It did, however, provide later examples of dative-fronted passives, up to the late fourteenth century. I concluded, based on the analysis of the passive that builds on targeting the immediately postverbal argument and the dative-fronted passive as a theme passive with a fronted dative, that the dative-fronted passive blocked the formation of a recipient passive with a nominal subject. When the dative-fronted passive was lost, this restriction no longer played a role and recipient passives became possible.

Section 3.4 investigated the introduction of the Exceptional Case Marking constructions, both active and passive, because the entire construction was new to the language, being introduced in the late fourteenth century. The discussion of Fischer (1989) and Los (2005) shows that the story of the ECM passive seems to be a tale of two separate constructions: on the one hand, the ‘true’ ECMs with *believe*-verbs and on the other hand, three-place verbs such as *command* that are reanalysed as two-place, i.e. ECM, verbs. The PPCME2 did not provide new evidence for this scenario of reanalysis with *command* verbs. For the *believe*-ECM, influenced by Latin (cf. Warner 1982), the data show that it was introduced as an entire construction, active *and* passive, meaning that the dominance of the passive cannot be explained with reference to the introduction (e.g. contra Noël 2008). Instead, I argued that it lies with the syntactic and surface similarity of the passive ECM to other passives that were already possible in English. This structural similarity made the passive more easily acceptable, while the active ECM continued to represent a marked structure.

The attribution of the development of the prepositional and recipient passive to individual changes in English explains the exceptional status of the passive in English within the family of (West)-Germanic languages: the changes can be ascribed to circumstances taking place in English which did not take place in languages like Dutch and German, such as the extended possibilities for stranding and the loss of dative fronting. Crucially, the changes do not rely on loss of case alone—something which has often been discussed in historical English work, but which overlooks the fact that Dutch lost case as well but did not develop in the same way as English. Returning to the overall question of the relation between the introduction of these passives and the loss of V2, I have shown that neither the timing nor the factors influencing the introductions indicate a direct relation. The new passives were introduced in the late fourteenth century, i.e. before what has usually been taken as the beginning of the loss of V2, 1400. The loss of V2 was not yet a factor which could drive a change like the introduction of the passives at the period that they were introduced—especially when what is most important, as in Los and Seoane’s proposal, are the information-structural effects of the loss of V2,

rather than the syntactic properties. Moreover, with the exception of the recipient passive, which is structurally related to the loss of V2 through the loss of the dative-fronted passive, there is no evidence of V2 playing a role structurally in the introduction of the passives. The following chapters assess in more detail the consequences of the loss of V2 and its role in the development of the passive throughout the history of English.

4. The Given-before-New Principle in Verb-Second Old English

Section 4.1 Introduction

This chapter investigates the information-structural properties of three Old English (OE) constructions: the passive, object fronting and PP preposing.⁵¹ Each of these constructions in OE involves movement and thus a reordering of arguments, which in all likelihood takes place for information-structural purposes. I focus on the two information-structural functions that were identified in Chapter 2 as relevant to the diachronic scenario for the passive and verb second (V2): the given-before-new principle and the use of the first position for local anchoring. The aim of this chapter is three-fold: (i) to establish to what extent these three constructions show an overlap in function in creating a given-before-new order; (ii) to establish to what extent the given-first, i.e. discourse-linking, part of the given-before-new principle is relevant to these constructions; (iii) to establish to what extent the syntactic movement in these constructions is driven by information-structural factors. I show that there is indeed an overlap in information-structural function between these three constructions, although only partial, and that even though they cannot in all

⁵¹ Some of the material in this chapter, in particular in Section 4.2.2 and 4.3.2, also appears in adapted form in Dreschler (2014).

cases be seen as the motivation for movement, the information-structural factors are clearly relevant to the use of these constructions.

Several authors have observed that the passive in Present-day English (PDE) does not have many functional alternatives. Siewierska (1984: 222) points out that PDE makes more use of the passive than related languages, and attributes this to the lack of other topicalizing and impersonalization strategies in PDE, where she defines topicalizing as placing given ('topical') information in clause-initial position. In related languages she finds an alternative for this function in object fronting as it is found in modern V2 languages, where the use of fronted objects is information-structurally unmarked—unlike in PDE. For PDE, Birner & Ward (1998) claim that the function of inversion, most common with prepositional phrases (PPs), is to create or restore a given-before-new order of information, similar to the passive (see Section 2.4.2). While some of these constructions are still available in PDE, the crucial difference lies in the fact that in PDE they are pragmatically marked.

These two constructions, object fronting and PP preposing, which could be alternatives for the given-before-new function of passives, are part of the V2 options in OE and crucially, there are no indications that their use was marked. While it is clear that object fronting is limited in PDE, OE had a similar type of unmarked object fronting to the one that is available in modern V2 languages (see Speyer 2010). At the same time, while it has been noted for modern English that the use of clause-initial PPs is marked (cf. Virtanen 1992; Biber et al. 1999: 803; Hasselgård 2010), unmarked discourse links are possible in this position in OE, as they are in modern V2 languages.

Within the V2 system in OE, then, it seems that there are alternatives for the passive in the more extensive range of word order options that are available for a speaker to create a given-before-new order of information. Indeed, recent studies have characterized V2 itself in terms of information-structural factors (Los 2009, Hinterhölzl & Petrova 2010). It has already been shown in the literature that word order in OE is at least to a certain extent determined by such factors (e.g. Bech 2001; Van Kemenade & Los 2006; Van Kemenade & Westergaard 2012; Taylor & Pintzuk 2012). What is not yet clear, however, is to what extent information-structural factors are the driving force behind specific constructions. Moreover, it has not been demonstrated that the passive already had a function of creating given-before-new order in OE to the same extent as Seoane (2000) has shown for Early Modern English or Birner (1996) for PDE, or to what extent constructions with similar information-structural motivations compete with each other.

In this chapter I argue that the passive, object fronting and PP preposing, crucially all involving reordering of arguments, were all available for OE speakers to achieve or maintain given-before-new order, while also having their own specific motivations. The main conclusions are that (i) the use of the long passive as an information-rearranger is less established than has been shown for later periods; (ii) both object fronting and PPs preposing are often used with a given-before-new order but also have a sizeable category of exceptions in marked, contrastive

readings; (iii) while all these constructions share the argument-reversal properties that allow for a reordering to achieve given-before-new order, the strength of the information-structural motivation for the movement is different for each construction.

The chapter is structured as follows. Section 4.2 describes the approach to the passive that will be used throughout the corpus studies in this chapter, following a discussion of the existing literature on the OE passive. It then presents a corpus study of two OE texts, *Orosius* and *Lives of Saints*, which finds that the passive is indeed used for given-before-new ordering, but that long passives are less frequent in the texts than they are in later periods. The corpus study also looks at the discourse-linking properties of passive subjects and assesses the information-structural properties of passive late subjects, concluding that while late subjects are used for placing new information late in the sentence, this is not their only function. Section 4.3 investigates the use of object fronting in the two selected texts and shows that the majority of object fronting sentences create a given-before-new order but that fronted objects can also be contrastive. It also shows that the majority of fronted objects function as ‘local anchors’, meaning that they provide a link to the immediately preceding sentence. Finally, this section’s corpus study investigates the information-structural character of pre- and postverbal subjects, the findings of which suggest a preference for given-before-new order, but not to the extent that information status directly relates to position. Section 4.4 discusses fronted PPs and shows that they, again, follow given-before-new order, but that there are two distinct additional categories. The study shows that the majority of clause-initial PPs, like objects, function as local anchors. It also shows that some of the inversion patterns of PDE are already present as distinctive patterns in OE. Section 4.5 concludes.

Section 4.2 The passive in Old English

This section investigates the use of the passive as an information-rearranger and the information-structural properties of passive subjects in OE. It aims to determine whether the passive was already used to create given-before-new order of information, and to investigate the information-structural properties of passives subjects in two contexts: (i) compared to active subjects; and (ii) in post-participle, or late position. In this section, I show that the long passive is already used as an information-rearranger in OE, but that it is as yet less established both as a syntactic construction and as an information-structural device.

Traditional works on OE syntax, like Visser (1963-1973) and Mitchell (1985), as well as studies focusing on the passive, such as Kilpiö (1989) and Denison (1993), present the following range of passives as found in OE.

- (1) *pær wæron gehælede þurh ða halgan femnan fela adlige men*
 there were healed through the blessed woman many sick men
 (ÆLS I 20.113; Denison 1993: 416, his 11)

- (2) *ac him næs getiðod ðære lytlan lisse*
but him(DAT) not-was granted that small favour(GEN)
'But he was not granted that small favour'
(*ÆCHom* I 23.330.29; Denison 1993: 108, his 29)
- (3) *Forðæm se ðe his ær tide ne tiolað, þonne bið his on tid untilad*
'then it will be unprovided in respect of him when the time comes'
(*Boethius* 67.11; Mitchell 1985: 355)
- (4) *and se munuc hatte abbo*
and that monk was-called Abbo
(*ÆLS* II 32.3; Denison 1993: 421, his 33)

Example (1) shows a sentence with the patient *fela adlige men* of the transitive verb *heal* as the subject of a passive sentence. Example (5) shows a passive with a fronted dative and example (3) with a fronted genitive. Next to the periphrastic passive with an auxiliary and past participle, as in examples (1)-(3), OE had a morphological passive with *hatan*, as in (6).

While it has been shown for EModE by Seoane (2000) and for PDE by Birner (1996) that the passive is subject to a given-before-new constraint, no such studies exist for OE. Since the long passive was already in use in OE, it might be expected that this is at least already one of the functions of the passive. However, there are additional impersonalization, information-rearranging and topicalizing strategies in OE that were lost in later periods, suggesting that the passive could have had a different function in OE than in PDE because there is more competition, or, alternatively, that it had not yet acquired all functions it performs in PDE.

This section is structured as follows. Section 4.2.1 reviews previous studies and studies of the OE passive, in an attempt to arrive at a clear definition of which types of sentences should be taken into account in a corpus study of the OE passive in *Orosius* and *Lives of Saints*. Section 4.2.2 presents the results of the corpus study of the information-structural properties of passives and focuses on three aspects of the OE passive: (i) the relative information status of subject and agent phrase in long passives; (ii) the information-structural behaviour of passive and active subjects; (iii) passive subjects that occur in a late position in the clause. Section 4.2.3 concludes.

4.2.1 A characterization of the Old English passive

This section aims to establish which factors of form and function are relevant to the selection of the passives in the corpus, and at the same time the discussion serves as a characterization of the passive in OE. There are several issues for the passive in this period that have been debated in the literature—most notably in Visser (1963-1973), Mitchell (1985), and Kilpiö (1989)—which are relevant to the corpus study of the passive in OE in the next section. What is most striking is that the passive in OE was not yet as standardized a construction as it is in PDE; on the other hand, some of

the issues concerning the classification of the passive in OE are familiar because they also were discussed for PDE in Section 2.3.

There are two issues relating to the form of the passive construction which are consequently directly relevant to the selection of passives in the corpora. The first is the widely debated variation between the two sets of auxiliaries that were available in OE to form a passive: example (5) below illustrates the use of *beon/wesan*, while example (6) illustrates the use of *weorðan*.

- (5) *Eala þu besceop to bysmore synd getawode þas earman landleoda.*
 behold thou bishop to shame are brought the poor-people of-this-land
 ‘Behold, thou bishop, the poor people of this land, are brought to shame’
 (ÆLS 32.64; Mitchell 1985: 305, translation from Skeat 1966)
- (6) *Se munuc þa Abbo... weorð sona to abode geset.*
 this monc then Abbo ... was immediately to abbot appointed
 ‘This monk Abbo ... was almost immediately appointed abbot.’
 (ÆLS 32.10; Mitchell 1985: 306, translation from Skeat 1966)

Weorðan is eventually lost as a passive auxiliary during the ME period (see e.g. Mustanoja 1960, Petré & Cuyckens 2008). It was already less common than *beon/wesan* in OE; Kilpiö (1989: 13, 61), for instance, finds no examples in Bede, and only 62 out of 641 passives in *Cura Pastoralis* are formed with *weorðan*. Mitchell (1985: 331) refers to Klingebiel (1937: 101-4) and Wattie (1930: 141), who find that *weorðan* is not common with the present tense, while it is “much more frequent” in the past tense. The debate regarding the nature of the variation between *beon/wesan* and *weorðan* has focused on the suggestion that the auxiliary expresses the stative (*beon/wesan*) or actional/change of state (*weorðan*) nature of the passive. Mitchell refers to Klæber (1923) and Frary (1929) for this distinction, but argues against it himself, claiming that *beon/wesan* could express “an action or process as well as a state” (1985: 330). Likewise, Kilpiö, even though he finds that *weorðan* is more common in the past tense and is only used in an “actional” way, calls Frary’s distinction “artificial” (1989: 64-67).

Petré (2010) revisits the idea of a functional distinction between *beon/wesan* and *weorðan*. He claims that in adjectival passives (*John was frightened*, cf. section 2.3.1) there is indeed a difference between the two, with *weorðan* indicating a change of state and *beon/wesan* an “existing state” (2010: 100). For the resultative passive (*The metal was flattened*, cf. Section 2.3.1), he proposes the following working definition of the difference: *weorðan* is “preferred for actions involving a high degree of energy exchange by both agent and patient”, while *beon/wesan* is “preferred for actions that lack such a high active involvement of the agent” (103). This is based on an examination of *weard/was ofslagen* examples from Orosius, ChronA and ChronE, first presented by Mitchell (1985: 332), who shows—contra Frary (1929)—that there is no distinction between *beon/wesan* and *weorðan*

ofslægen. Mitchell shows that of 186 examples with *ofslægen*, based on data from Hoffman (1934: 12), 79 are used with *wæs/wæron* and 107 with *weorð/wurdon*, and concludes that they are used interchangeably. This is where Petré brings in the difference in context: “Whereas *weorðan* is invariably used when the patient was killed in battle *while fighting*, and the expression is part of the main narrative, *wæs* is used in all other cases, all (except for the backgrounded ones) involving some form of dying without fighting (mostly murder)” (2010: 105). Petré then suggests, based on data from the late chronicle A&F, that the distinction is “bleached and probably finally lost altogether during late Old English” (*ibid.* 106). While the potential difference between the two auxiliaries is interesting, it is clear that they cannot be used straightforwardly to distinguish between different types of passives, and for that reason both sets of auxiliaries will be included in the corpus studies in this chapter.

The second difference in form with PDE is that there is no standardized way yet to express the agent in OE. Mitchell (1985) and Kilpiö (1989) point out that the agent can also be expressed as an instrument with a case form, as in example (7).

- (7) *þa þe manna handum geworhte wæron of eorðlicum timbre, ...*
which by men’s hands formed were from earthly timber
(Bede III 2345 (O); Kilpiö 1989: 154, translation mine)

However, Mitchell states that the expression of an agent with a prepositional phrase is “the norm” (1985: 335). He presents the range of prepositions that can be used for this purpose and discusses these in quite some detail: *æt, be, for, fram, mid, of, þurh, wið* (1985: 336-348). The most interesting of these is of course *be*, the predecessor of present-day *by*, although it is by no means the most frequently used preposition in OE, and often still has a different meaning; Mitchell claims that *be* was not used to express the “principal agent” (337). He also states, however, that there are some examples with *be* which are similar to the PDE use of *be*: these “come close to the first use” as a principal agent, rather than what he calls the immediate or subordinate agent (337). He gives the following example:

- (8) *Sum eawfæst man sende ðam halgan were twegen butrucas mid wine to lace,*
be anum cnapan.
‘A pious person sent to the holy man two flasks of wine as a gift, by a boy.’
(ÆCHom ii. 170. 13; Mitchell 1985: 337, translation from Thorpe 1846 [2013]: 171)

While Mitchell states that this example is close to the PDE use, he also points out the possibility that it is not just the lack of a fixed preposition but also the lack of the notion of a personal agent that distinguishes the OE passive from later periods: he proposes that it is possible that “the personal agent in the sense in which we mean it today was not expressed in English until the usage with ‘by’ became established”

(1985: 347). In other words, while we find long passives in OE texts, they are not straightforwardly comparable to PDE long passives.

Kilpiö (1989), like Mitchell (1985), discusses the difficulty of determining what constitutes a real agent and ultimately defines the agent as “originator or the instrument of the action expressed by the passive predicate, and which in most instances can be converted into the subject of the corresponding active construction” (1989: 136). Kilpiö lists fewer prepositions than Mitchell and discusses these on the basis of frequency, with *fram* (9), *þurh* (10), and *mid* (11) as the most common ones.

(9) *þa ðe wæron fram nædran geslegene*

then they were by an-adder slayn

(Bede I 139 (Ca); Kilpiö 1989: 140, translation mine)

(10) & *ðurh us scylen bion hiora scylda gestiered mid cræfte & mid lare.*

and through us shall be their shields guided with strength and with knowledge

(CP(H) 117.20; Kilpiö 1989: 144, translation mine)

(11) & *sona swa hit bið forlæten from Gode, swa bið hit gedrefed mid diofles ege.*

and immediately when it will-be left by God, so will-be it disturbed by the-devil's eyes

(CP(H) 465.13; Kilpiö 1989: 147, translation mine)

In other words, both Mitchell and Kilpiö show that there is no straightforward or syntactically fixed way to form long passives in OE. The consequence of this as yet non-standardized way to express the agent is that it is not possible to select agent phrases in the corpus on the basis of form only.

In addition to these two issues of form, there have been attempts to classify the OE passive semantically or functionally, similar to the distinctions that have been made for PDE. There are two types of distinctions that are generally made: stative and dynamic on one hand, and adjectival, resultative and verbal passives on the other hand. Toyota (2008) presents a study of the distinction between stative and dynamic passives throughout the history of English, starting with OE. He dismisses the traditional tests for stativity (most importantly the choice of auxiliary, but also the use of the *ge-* prefix on the participle which has been said to mark stativity), and bases his classification primarily on “undergoer orientation” and the “presence of outer cause” (24), which according to Toyota both point to a dynamic passive. His results show an increase in the proportion of dynamic passives during the OE period. Table 4.1 repeats his data, showing the percentage of dynamic, stative and ambiguous passives for each of the four OE subperiods.

Table 4.1 Aspectual change during OE (Toyota 2008: 17, his 2.3)

	OE1	OE2	O3	OE4	Total
Dynamic	0 (0%)	43 (23.3%)	157 (32.1%)	195 (40.7%)	395 (34.2%)
Stative	2 (100%)	122 (65.9%)	293 (59.9%)	244 (50.9%)	661 (57.2%)
Ambiguous	0 (0%)	20 (10.8%)	39 (8.0%)	40 (8.4%)	99 (8.6%)
Total	2 (100%)	185 (100%)	489 (100%)	479 (100%)	1155 (100%)

Toyota's data for OE already show a considerable change between the periods in the percentage of stative passives: from 65.9% in OE2, to 50.9% in OE4. That said, the tests for stativity are not entirely reliable and it remains to be seen how likely it is that such a development already takes place throughout the OE period, considering the differences in corpus size per subperiod. Toyota's study does suggest that the distinction between stative and dynamic is important in the history of the English passive, because stativity seems to be an area of change even in OE. At the same time, however, the distinction between the two categories cannot be made on the basis of form, and, as discussed above, there are few reliable clues: Toyota shows that many of the tests—such as the selection of auxiliaries—are not reliable, but he himself also uses notions which are ultimately multi-interpretable, such as 'presence of outer cause'.

Toyota (2008) and Petré (2010: 37-40) distinguish the same three types of passives that were discussed in Section 2.3.1 for PDE passives: adjectival, resultative and verbal. Examples (12)-0 from Toyota illustrate these three categories:

- (12) *Forþon ic eom gesett betweenen þisum folce, swa swa sceap betweenon wulfum and ic eam befangan eal swa spearwe on nette ...*
 since I am set between these folks just like sheep between wolves and I am enclosed completely like sparrow in net
 'since I am trapped between these people just like a sheep between wolves and I am completely enclosed like a sparrow in a net...'
 (HC OE4 comarga; Toyota 2008: 30, his 32)
- (13) *lc secge eow to soþan þæt sib is forgifen Godes gelaðunge*
 I say to you in truth than peace is given God.gen congregation
 'I say to you truly that peace is granted to God's congregation.'
 (HC OE3 coelive; Toyota 2008: 29, his 26)

- (14) *& he his feorh generede & beah he wæs oft gewundad*
 and he his life saved and yet het was often wounded
 ‘and he saved his life, although he was often wounded’
 (HC OE2 cochroa2; Toyota 2008: 31, his 37)

Petré points out, however, that the distinctions between the three types are sometimes difficult to determine. One aspect that he presents as aiding identification of adjectival and resultative passives is the fact that predicative participles, like predicative adjectives, “generally agree in case and number with the Subject”, which he illustrates with (15).

- (15) *And þa Iudeas wæron ofslagene þam fulostan deaðe and heora naman syndon adylegode ofer ealre eorðan.*
 And the Jews were killed:NOM.PL the foulest death and their names:NOM.PL are destroyed:NOM.PL over whole earth
 ‘And the Jews were killed with the foulest death and their names are erased over the whole earth.’
 c1075 (VSaI 1 [Cross]: 30.4; Petré 2010: 40, his 55)

Petré points out that, in contrast to case and number, there is no agreement in gender, but he refers to Mitchell’s (1985: 16-21) observation that such agreement is not always present for predicative adjectives either. This observation about form means that it is, to a small extent, possible to make the distinction between verbal, adjectival and resultative passives on the basis of form alone, i.e. the presence of agreement.

Petré (2010) discusses the overlap between adjectival and resultative passives and what he calls adjectival copular constructions: intransitive structures with a non-agent subject. The verbal passive, in contrast, is not intransitive, because “the Subject is no longer merely a non-agent, but is the patient of a transitive event, whose personal agent is also always evoked by the context (if not necessarily expressed)” (2010: 40). The difference between adjectival/resultative and verbal, in other words, does not only lie with the more often dynamic nature of the verbal passive, but with the implication of an agent (cf. Toyota’s distinctions). Petré suggests that in OE the verbal passive had not yet developed from an earlier adjectival use, referring to Mustanoja (1960: 440), Toyota (2008), McFadden & Alexiadou (2006), and Haspelmath (1992) for (cross-linguistic) studies which show that adjectival use precedes the verbal use. Petré’s discussion shows that there is an (inherent) overlap between different types of passives and that these distinctions cannot be made on the basis of form alone. Moreover, as discussed in Section 2.3.1, these distinctions may not be essential for the present study because of the focus on information-structural aspects, which are relevant for both verbal and adjectival passives. Therefore, these distinctions will not be included in the corpus studies.

Yet another issue that is specific to OE is the fact that there is a potential overlap with perfect constructions. This overlap, or ambiguity, arises from two different factors. The first factor is that while the perfect in PDE is always formed with *have* as the auxiliary, OE also uses *beon/wesan* and even *weorðan*, the same auxiliaries that are used for the passive. This overlap is only small, because, as Mitchell and Petré both state (on the basis of different corpora), this type of perfect is not frequent, and in fact, Mitchell (1985: 299-304) questions whether *beon/wesan/weorðan* are real auxiliaries in these perfects, instead of presumably copulas. The second factor, already discussed in Section 3.2.3.2, is that there was no marking of perfectivity on passives, i.e. *the book is written* could both be a present passive as well as a perfect passive *has been written*. This means that a passive of a present perfect and a simple present look similar, witness (16), repeated from Section 3.2.3.2.

(16) *þa wæs an gereord on eallum mancynne. and þæt weorc wæs begunnen ongean godes willan;*

then was one language among all mankind and the work was begun against god's will

'At that time all people spoke one language, and the work had been begun against God's will'

(ÆCHom I, 22 318.17; Los 1999: 103, her 11)

Recall from the discussion in chapter 3 that there were two possible interpretations for these causative/ergative pairs, as pointed out by Los (1999: 103): either a passive of 'X began the work' or the perfect 'the work had begun'.

I have discussed in this section the most important aspects of form and selection of passives in OE: the choice of auxiliary between *beon/wesan* and *weorðan*, the variety of ways to express the agent, the semantic classification such as the distinction between adjectival and verbal passives and stative vs dynamic passives, and finally, the overlap between passive and perfect. While some of these distinctions are in theory relevant to any study of the passive in OE, I have argued that the choice of auxiliary and the functional classification are not crucial for the present study because the focus is on information-structural factors, which are not restricted to one type of passive. I therefore conclude that the selection of passives should be determined solely on the combination of one of the auxiliaries *beon/wesan/weorðan* and a past participle. While using this method of selection, it should be taken into account that there is the probability that some of the examples are in fact copular/adjectival constructions. In contrast, the selection of long passives should be done manually because of the different prepositions that can express the agent.

4.2.2 A corpus study of passives in YCOE

This section presents the results of a corpus study of the passive in two OE texts, *Lives of Saints* and *Orosius*, from the YCOE. I show that long passives are indeed used for information-rearranging purposes. I also show that there are no clear indications that passive subjects behave differently from active subjects. Finally, I argue that passives are relevant to late subject constructions because they provide a grammatical option for a subject to surface late, like unaccusatives, which can be used for information-structural purposes.

4.2.2.1 Method of selection and analysis

Selected texts

All studies in this chapter are based on *Orosius*, a text in the West Saxon dialect from the second subperiod of the YCOE (O2, 850-950), and *Lives of Saints*, a slightly later text (from O3, 950-1050) in the same dialect.⁵² The two texts are among the longest in the corpus, and as such offer enough material for the specific constructions under investigation in this chapter. *Lives of Saints*, classified in the YCOE as ‘Biography’ and ‘Lives’, is a collection of saints’ lives—i.e. short narratives—but also contains several sections that are more expository in nature. *Lives of Saints* is written by Aelfric, who is known as a highly stylistic writer, which may in some cases lead to idiosyncrasies in the text (see e.g. Sato 2012, Ohkado 2004). *Orosius* is classified by the YCOE as ‘History’ and contains descriptions and narratives of historical events. At the same time, it also contains parts that are similar to a travelogue in that these describe regions and locations of countries with respect to each other or other geographical features. Most importantly, both texts contain large stretches of narrative, which makes them particularly suited for a study of information-structural factors: there are many instances of new referents that are introduced in the discourse, while other referents remain the topic of the conversation for a considerable length of time.

Selected clauses

Because the aim of this section is to investigate the information-structural properties of the passive as a rearranging device and so-called subject creator, rather than defining the OE passive in terms of semantic or functional character, the selection of the passive is based on form only: any main clause with a form of *to be* and a past participle is selected. The YCOE does not have a separate label for passive participles, but through the added selection of *be* as an auxiliary, the majority of the unambiguous perfective past participles (i.e. with *have*) can be excluded. The OE

⁵² I used the translation from the edition used for the YCOE for *Lives of Saints* (Skeat 1966), and the translation from Bosworth (1859) for *Orosius*, because the edition used for the YCOE (Batley 1980) does not contain a translation. In some cases, I have modernized the translation or changed it to represent the order of elements in the OE examples.

auxiliary *weorðan* is coded as *be* in the YCOE, and is thus included automatically in the following query.

<i>Query 4.1</i> Passive main clauses in the YCOE
a. IP-MAT idoms *BED* *BEP*
b. IP-MAT idoms *VBN*

Line 4.1a selects all main clauses with a form of *to be*, while line 4.1b selects all main clauses with a past participle as constituent. Only main clauses were selected because of the focus in this thesis on V2, a main clause phenomenon. Selecting only main clauses will allow for a straightforward comparison of passives with the V2 alternatives for information rearranging, object fronting and PP preposing. To keep the database as clean as possible, I only included IP-MAT main clauses, instead of IP-MAT*, meaning that direct speech is excluded, as well as appositive or parenthetical sentences. Excluding direct speech is especially relevant from the perspective of information-structural coding, because for direct speech the context in which an utterance is expressed is even more difficult to determine. For those aspects of the studies where the position of the subject was of interest, I included one extra element in the query, namely that IP-MAT should immediately dominate ('idoms') a non-ellipted subject.

For the selection of long passives I added the following element to the query:

<i>Query 4.2</i> Prepositional phrases
a. IP-MAT idoms PP*

This query was executed on the results of Query 4.1, yielding all passive clauses which also contain a prepositional phrase; these sentences were manually analysed to determine whether the prepositional phrases were indeed agent phrases. Again, because the focus lies with information-structural aspects rather than grammatical or semantic aspects, the selection of agent phrases was not restricted to personal agents, and some PPs denoting instruments were also included, thus increasing the number of long passives in the selection.

Method of analysis

The subject and agent phrases of the passives in OE were analysed with respect to three aspects: information status, anaphoricity, and persistence. For information status, four categories were used, which are mainly based on Prince's (1981) work, and which build on existing annotation practices for OE as described in Section 2.4.3. The first category is OLD, which consists of discourse-old referents, illustrated in (17) and (18). The examples provide the context in which the sentences occur, with the relevant antecedents underlined.

- (17) ‘The soul is a rational spirit [...]. The benevolent Creator let it have the command over its own choice; then it became corrupt by its own free-will, through the Devil’s teaching;’
Heo wærð æft alysad þuruh Godes gife, gif heo Gode gehyrsumað.
 it shall-be again delivered through God’s grace, if it God obeys
 (coaelive [Christmas]:176.135)
- (18) ‘...and they led Abraham’s brother’s son Lot, with his family, out of the city; [...]’
and Loth se rihtwisa wearð ahred ðurh God.
 ‘and Lot, the righteous, was delivered by God’
 (coaelive [Pr_Moses]:211.2980)

As can be seen from the provided context, the subjects in (17) and (18), *Heo* and *Loth se rihtwisa*, have been mentioned before. The antecedent for *Heo* in example (17) is in the immediately preceding clause, while the antecedent for *Loth se rihtwisa* is three clauses preceding the one in which it is mentioned. This difference in distance to the antecedent is not relevant to the information status category, but is captured by the additional measure of anaphoricity.

The second category, ACCESSIBLE, contains elements which have not been present in the discourse, but which are nonetheless considered to be “shared knowledge” of both speaker and hearer (cf. Prince 1981: 230). This category is, in part, similar to Lambrecht’s (1994) notion of accessible (hence the term), but with the difference that Lambrecht uses the term ‘accessible’ for referents that *have* already been mentioned in the discourse—his definition is that they are present in the interlocutor’s “peripheral consciousness” but are not “active” anymore (1994: 76). The difference between the OLD and ACCESSIBLE categories as presented here lies in the fact that OLD referents are necessarily mentioned in the preceding discourse, while ACCESSIBLE referents are not. Examples (19) and (20) illustrate.

- (19) *and he wearð gehæled sona þurh þone ælmihtigan God, for Swyðunes geearnungum,*
 ‘and he was healed instantly by the Almighty God, for Swithun’s merits’
 (coaelive [Swithun]:394.4466)
- (20) *Ðas martyras næron næfre on life þurh wifbesmytene,*
 these martyrs not-were never in lives through women defiled
 ‘These martyrs were never, throughout their lives, defiled with women’
 (coaelive [Eugenia]:380.420)

Both relevant referents in these examples, *the Almighty God* and *wif*, have not been mentioned in the preceding discourse, but they can still be considered as shared knowledge: *God* is part of the speaker’s and hearer’s world knowledge in the context

of the OE saints' lives, while *women* is generic and thus accessible. The category ACCESSIBLE also includes elaborating inferrables, i.e. roughly speaking, referents which can be inferred from the mention of a related concept. For instance, if *bus* has been mentioned, *driver* is inferrable through the knowledge that buses have drivers (see Birner 2006, Taylor & Pintzuk 2011, 2014).

The third category, ANCHORED, is used for referents that are new to the discourse, but at the same time are linked to the discourse through an 'anchor' which is discourse-old.

- (21) 'Elisha also healed a nobleman from the dreadful leprosy; he was called Naaman, of the land of Syria, and he believed in God through the great miracle which God wrought in him.

Fela oðre tacna gefremede God þurh hine on Israhela ðeode,

'Many other signs performed God through him among the nation of Israel'

(coelive [Book_of_Kings]:313.3892)

- (22) 'The sea, Propontis, lies on the east of Constantinople.'

& be norðan Constantinopolim Creca byrig scyt se sæearm up of þæm sæ wstrihte þe man hæst Euxinus;

and on the.north.of.Constantinople.of.the.Greeks city shoots the sea-arm up from the sea western that they call Euxine

'and on the north of Constantinople, a city of the Greeks, the arm of the sea shoots up right west from the Euxine;'

(coorosiu,Or_1:1.18.5.349)

Both *Fela oðre tacna* and *be norðan Constantinopolim* are new to the discourse but contain an element which is linked to the discourse and which in turn connects these subjects to the preceding discourse. The ANCHORED category also includes Birner's (2006) second type of inferables—bridging inferables—which are inferable elements which are not accessible without the anchor, in OE often expressing alienable possession (see Taylor & Pintzuk 2011, 2014).

The fourth and final category is that of NEW referents, i.e. referents that are newly introduced to the discourse, without any element that links them to the discourse, as illustrated in (23) and (24).

- (23) *Sume gedwolvmen wæron þuruh deoful beswicane*

'Certain heretics were by the devil beguiled'

(coelive [Christmas]:7.6)

- (24) *þær wæron gehælede þurh ða halgan femnan fela adlige menn,*

'There were healed by the holy woman many sick men,'

(coelive [Æthelthryth]:113.4208)

Neither *Sume gedwolmenn* nor *fela adlige menn* are in any way inferable or anchored to the discourse. The category includes all types of new referents that were discussed in Section 2.4.3.2: specific and non-specific referents, inert referents, and short-term referents, i.e. referents which are available only within a limited context (see Taylor & Pintzuk 2011, 2014 and Haug et al. 2009, 2014). Table 2.5 summarizes the information status categories, including some of the subcategories from the literature.

Table 4.2 Annotation scheme for information status

Category	Categories from other models included in each category	Binary distinction
OLD	Prince's (1981) <i>Textually Evoked</i> , Komen et al.'s (2014) <i>Identity</i> , Pintzuk & Taylor's (2011, 2014) <i>Previously mentioned</i> .	'Given'
ACCESSIBLE	Prince's <i>Unused</i> and <i>Situationally Evoked</i> , Komen et al.'s <i>Assumed</i> , Taylor & Pintzuk's <i>Shared/Cultural knowledge</i> , <i>Situationally Evoked</i> , Haug et al.'s (2014) <i>Generic</i> , Birner's (2006) <i>Elaborating inferables</i>	
ANCHORED	Birner's <i>Bridging inferables</i> , Prince's <i>Brand-new Anchored</i>	'New'
NEW	Prince's <i>Brand-New Unanchored</i> , Taylor & Pintzuk's and Haug et al.'s <i>Specific</i> , <i>Non-specific</i> and <i>Short-term referents</i> , Komen et al.'s <i>Inert</i>	

Note that these distinctions are preliminarily considered to be a scale or hierarchy, implying that ACCESSIBLE referents are *more given* than ANCHORED referents, although the evidence in the corpus studies may lead to a different conclusion. Any ambiguous cases will be discussed where relevant in the discussion of the corpus studies. The first two categories constitute what is usually considered 'old' or 'given' in a binary distinction, and the last two categories constitute what is generally taken as 'new'. Where applicable, I will use a binary distinction, merging the categories, to compare my findings to previous studies which have used such a binary distinction.

The other categories of analysis—anaphoricity and persistence—calculate the presence in the discourse within a selected window defined as 5 preceding and 5 following main clauses of the relevant main clause, calculated automatically by *CorpusStudio*. These measures are based on Gregory & Michaelis's (2001) study on the comparison between topicalization and left-dislocation, in turn based on Givón (1983, 1984). For persistence, Givón originally considered 10 subsequent clauses, which Gregory & Michaelis bring back to 5 clauses. Based on earlier work, which showed that most anaphoric links were either close or non-present (Dreschler 2010), and the current interest in *local* links (Los & Dreschler 2012), the window for anaphoricity was also limited to 5 clauses. This smaller window for anaphoricity also allows for a distinction between anaphoricity and information status: if all or most of

the preceding text is taken into account discourse-old referents are always also anaphoric. Using a smaller window of 5 clauses means it is possible to distinguish between a referent that has been mentioned 20 clauses before (no previous mention but OLD) and 1 clause before (previous mention and OLD). Persistence in Givón's (1983, 1984) and Gregory & Michaelis's (2001) work was used specifically for aboutness-topic persistence, but is used here for any referent, even when it is not the topic. Table 4.3 shows the categories.

Table 4.3 Annotation categories for anaphoricity and persistence

Anaphoricity	Persistence
Distance to antecedent: 1-5	Distance to subsequent mention: 1-5
Distance to anchor: 1-5	Distance to subsequent inference: 1-5
No prior mention	No subsequent mention

The antecedent, anchor, subsequent mention and subsequent inference can be either a subject or a complement (object or part of a prepositional phrase). Note that the table distinguishes between actual prior mention and the reference to an anchor; this also includes reference to a set of which one member is selected, as in *Three categories...The first category...*

4.2.2.2 The order of information in long passives

This section presents a study of the relative information status of the subject and the agent phrase in long passives, in an attempt to establish whether the observations with regard to the information-rearranging function of the passive in PDE by Birner (1996) and in Early Modern English by Seoane (2000) already apply to the OE passive. The data show that there are only a handful of exceptions to the given-before-new order of information and so the majority of examples indeed have an information-rearranging function, restoring the unmarked given-before-new order of information in the sentence.

Table 4.4 first shows the frequencies of both types of passives in the OE texts used for this chapter.

Table 4.4 Long passives in main clauses in *Orosius* and *Lives of Saints*

	Short passive		Long passive		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
<i>Orosius</i>	191	84.5	35	15.5	226	100.0
<i>Lives of Saints</i>	465	92.6	37	7.4	502	100.0
Total	656	90.1	72	9.9	728	100.0

Between the two texts: $\chi^2(1) = 11.52$, $p < .001$

As can be seen in the table, there is a considerable difference between the two texts in the percentage of long passives: 15.5% long passives with respect to short passives for *Orosius*, against 7.4% for *Lives of Saints*. Additionally, there is a considerable difference between the average for these two OE texts (9.9%) and the percentage found by Seoane (2000), as well as percentages for Modern English in several studies. Toyota (2008: 11) claims that 20-30 percent of the passives in PDE are long, referring to Jespersen (1924: 168), Svartvik (1966: 141), Givón (1979: 57-64), Huddleston (1984: 441) and Dixon (1991: 278). Seoane's (2000) percentages for EModE are somewhere in-between the percentages for PDE and the data from *Orosius* and *Lives of Saints*: 15.4% on average for EModE (2000: 28). The lower frequencies add to the picture of the passive as a construction that was not clearly defined yet in all respects.

Table 4.5 presents the results of the analysis of the relative information status of subject and agent phrase. Note that the total number of long passives included in this table is 63 and not the 72 given in the table above: only long passives with a subject preceding the agent phrase and with the agent in a prepositional phrase were included. The other long passives are discussed separately below. The information status of the subject is given in the rows and the information status of the agent in the columns. The shaded cells indicate those instances in which the information status of the subject is lower than or equal to that of the agent phrase.

Table 4.5 Relative information status of subject and agent phrase in long passives

		Agent phrase				TOTAL
		OLD	ACCESSIBLE	ANCHORED	NEW	
Subject	OLD	41.3% (26)	15.9% (10)	9.5% (6)	22.2% (14)	88.9% (56)
	ACCESSIBLE					
	ANCHORED		3.0% (2)			3.2% (2)
	NEW	1.6% (1)	3.2% (2)	3.2% (2)		7.9% (5)
	Total	42.9% (27)	22.2% (14)	12.7% (8)	22.2% (14)	100.0% (63)

$\chi^2 (6) = 13.38, p < .05$

Table 4.5 shows that a majority of the sentences, 56 of 63 (88.9%), fall into the shaded cells: the information status of the subject is equal to (26 examples, 41.3%) or lower than (30 examples, 47.6%) the information status of the following agent phrase. In other words, the majority of examples follow the given-before-new principle. The percentage of 88.9% is slightly lower but still comparable to Seoane's (2000) figures: she found that 95.9% of the examples either showed given-before-new order or an equal information status of subject and agent phrase. Finally, the information status of the agent phrase is clearly different from that of the subject: only 42.9% of the agent phrases are OLD, against 88.9% of the subjects, and 22.2% of the agent phrases are NEW, against only 7.9% of the subjects.

Examples (25)-(27) illustrate the categories where the information status of the subject is more given than that of the agent phrase. In these and following examples, the subjects are given in bold and the agent phrases are underlined.

(25) =(20)

Ðas martyras næron næfre on life burh wif besmytene,
‘These martyrs were never, throughout their lives, defiled with women’
(coelive [Eugenia]:380.420)

(26) *Pa wearð **Alexander** ofslagen his broðor from his agenre meder*

then was Alexander slain, his brother, by his own mother
(coorosiu,Or_3:7.61.15.1175)

(27) *ac **heo** wæs eall freten ... burh fule hundas.*

‘but she was all devoured ... by foul dogs’
(coelive [Book_of_Kings]:352.3921)

In all three examples, the subject is OLD—*Ðas martyras* (25), *Alexander* (26) and *heo* (27)—while the agent phrases belong to different categories: ACCESSIBLE *wif* in (25), ANCHORED *his agenre meder* in (26), and NEW *fule hundas* in (27). Example (28) below illustrates the category where the subject and agent phrase have the same information status, OLD.

(28) *and **hi** wurdon gehælede, burh þone halgan wer.*

‘and they were healed through the holy man’
(coelive [Oswald]:200.5497)

In this example, both *hi* and *þone halgan wer* have been mentioned in the discourse before and are therefore both OLD.

As for the 7 exceptions to the expected pattern, i.e. instances where the agent phrase is more given than the subject, there are a number of observations worth making. First of all, there seem to be at least 4 examples which can be characterized as “event-reporting” in Lambrecht’s terms (1994: 124). These sentences, according to Lambrecht, have a different information structuring pattern (i.e. no regular topic-comment structure), which could explain why they are exceptions to given-before-new in this database. Examples (29) and (30) illustrate, both with a NEW subject and an ACCESSIBLE agent phrase.

(29) =(23)

Sume gedwolmenn wæron buruh deoful beswicane
‘Certain heretics were by the devil beguiled’
(coelive [Christmas]:7.6)

- (30) *þær wurdon gehælede æt ðære halgan byrgene eahta untrume menn, ... wundorlice þurh God.*
 ‘There were healed, at the holy tomb, eight sick men, miraculously, by God’
 (coelive [Swithun]:132.4296)

Example (29) is one of the first sentences in this particular Saints’ life, and both the subject *Sume gedwolmenn* and the agent phrase *þuruh deoful* are new to the discourse. In example (30), the focus is on the entire event, rather than one particular aboutness topic, while the link to the previous discourse is formed through the prepositional phrase *æt ðære halgan byrgene* rather than through the subject.

Another aspect that stands out in the exceptions to the given-before-new pattern is that many of the subjects occur after the participle, i.e. they are late subjects (see Warner 2007 for a discussion of late subjects in ME). Example (30) above and (31) and (32) below illustrate.

- (31) *Oft wurdon eac gehælede fela untrume men þurh his reafes fnæda*
 ‘Often were likewise healed many sick men by the borders of his garment’
 (coelive [Martin]:569.6335)

- (32) *On ðere stowe beoð gehælede gehwilce untrume þurh þæra martyra gearnunga þe on þære stowe ðrowodon.*
 on that spot are healed any sick-people [that come] through the martyrs’ merits who in that place suffered
 (coelive [Julian_and_Basilissa]:125.1012)

Late subjects are known to have a separate information-structural motivation and to favour new(er) subjects, which could indicate that in the late subject constructions this motivation is more important than retaining a given-before-new order of subject and agent phrase. It also means that there are other elements occurring before the subject (*þær* in (30) *Oft* in (31) and *On ðere stowe* in (32)), whose information status may be equally relevant. Note also that the agents in (31) and (32) are not only more given, but are also both much heavier than the subject: the subjects are *fela untrume men* and *gehwilce untrume* (both NEW), while the agents are postmodified with a relative clause, the ANCHORED *þurh his reafes fnæda*, *þe fela men of atugon*, and *bundon on þa seocan* in (31) and the OLD *þurh þæra martyra gearnunga þe on þære stowe ðrowodon* in (32). The agent phrase in (30), however, is light, only consisting of one word, the ACCESSIBLE *God*.

Finally, one of the exceptions to the given-before-new order—example (33), showing ANCHORED before OLD—suggests that it is not just the the information status that determines the order of elements, but that the presence of and distance to an antecedent plays a role as well (cf. Birner & Ward’s (1998) observation about passives, Section 2.3.2).

- (33) *and **eall his team** wearð gewurðod þurh God.*
 ‘and his whole family was honoured by God’
 (coaelive [Æthelthryth]:5.4147)

The subject *eall his team* is linked to the previous sentence through the anchor *his*, while the agent *God*, though discourse-old, has not been mentioned in the immediately preceding five sentences. This suggests that anaphoricity and information status are both relevant, but should be considered as independent factors. The same reasoning might be applied to the following example, where the subject is linked to the previous sentence by an anchor, while the agent has a weaker, inferred, link.

- (34) *Æfter þam, an þunor tosloh heora Capitolium, [þæt hus] þe heora godas inne wæaron, & **hiora biblioteoco** wearð onbærned from ligette,*
 afterwards a thunderbolt shattered their Capitol, the house which their gods in were, and their library was set-on-fire by lightning
 (coorosiu,Or_6:14.142.1.2976)

This example is interesting for yet another reason, because *biblioteoco* and *ligette* represent two different types of inferables, with one having an explicit contrast to an antecedent *þunor* and the other being assumed to be part of a city, which has been mentioned four sentences before this sentence. The discussion of these exceptions shows that there is no single explanation to account for the fact that these examples do not show the expected given-before-new order, but there are known factors, such as anaphoricity and event-reporting sentences, that go some way toward explaining why such exceptions occur.

Table 4.5 above only included those long passives where the subject precedes the agent phrase. There are four clauses in the database in which the agent phrase occurs before the subject. These sentences, while having an unexpected order of arguments, do nevertheless show the expected given-before-new order of information, as illustrated for two of these examples in (35) and (36).

- (35) *Þær wurdon gehælede þurh ðone halgan wer **feower wanhale menn** innan ðrym dagum,*
 ‘There were healed, by the holy man, four sick men within three days’
 (coaelive [Swithun]:143.4303)

- (36) *Mid þæm bryne **hio** wæs swa swiþe forhiened þæt hio næfre sibþan swelc næs,*
 by that burning it was so much wasted that it never afterwards such not-was
 ‘It was so much wasted by that burning that it never afterwards was such [as it had been]’
 (coorosiu,Or_6:1.133.15.2815)

Note that the first sentence is similar to examples (30), (31) and (32) in that the subject occurs after the participle. While the order of subject and agent phrase is reversed, the order of given-before-new is maintained: in (35), *þurh ðone halgan wer* is OLD, while *feower wanhale menn* is NEW, and in (36) both *Mid þæm bryne* and *hio* are OLD. In (36), the PP clearly functions as a local anchor in providing a link to the preceding discourse, one of the main functions of presubject material in OE (cf. Los & Dreschler 2012).

There are 4 examples in *Lives of Saints* (and none in *Orosius*) which have what is translated as an agent phrase, but where the agent is expressed through a case-marked NP (cf. Mitchell's example in (7)).

- (37) *and seo hlæfdige bið þære wylne underðeodd;*
and the mistress is the bondmaid.ACC subjected
'and the mistress is subjected to the bondmaid'
(coelive [Auguries]:10.3541)
- (38) *Heo wæs mannum gepuht swylce heo his gemæcca wære.*
she was men.DAT thought as-if she his wife were
'It seemed to men as if she were his wife.'
(coelive [Basil]:470.787)
- (39) *Ðas feower godspelleras syndon Gode gecorene,*
These four Evangelists are God.DAT chosen
'These four Evangelists are chosen by God,'
(coelive [Mark]:174.3309)
- (40) *ac Dunstan and Abelwold wæron Drihtne gecorene,*
but Dunstan and Æthelwold were God.DAT chosen
'but Dunstan and Æthelwold were chosen by God,'
(coelive [Swithun]:457.4499)

Example (37) has an OLD subject and an OLD agent; examples (38)-(40) have an OLD subject and an ACCESSIBLE agent. These examples were not included in the table because of their different syntax, but interestingly, they do follow a given-before-new order of information: the information status of the subject (mentioned first) is more given than or equal to the information status of the agent NP.

This study of long passives in *Orosius* and *Lives of Saints* has shown that the majority of long passives, roughly 90%, show an order of information which conforms to the given-before-new order of information for subject and agent phrase. These data, similar to the findings for PDE and EModE, provide evidence that the long passive was already used as an information-rearranger in OE. At the same time, the percentage of long passives in the two selected texts in comparison to short passives is lower than in later periods, suggesting that the (long) passive was

perhaps as yet less important as an information-rearranging strategy in OE. This correlates with the findings in Section 4.2.1 that the syntactic character of the long passive is also less stable at this time, with different ways to express the agent and more options to order agent phrase and subject.

4.2.2.3 Passive subjects as discourse linkers

In Chapter 2, we already saw that it is not just the relative information status of subject and agent phrase that is relevant to the use of the passive, but part of the scenario presented by Seoane (2006) and Los (2005, 2009) that explains the rise of the passive is based on the discourse-linking function or topicality of the passive subject. The focus in this section is on the information status of passive subjects, rather than for instance topicality, and for two reasons. Firstly, although Seoane's scenario is based on topicality, her data show that of all features contributing to topic status, information status (a binary given/new in her studies) is the most relevant factor. Secondly, the information status of the subject is the central notion in Los's scenario of unmarked themes: "old information with [a] lack of prominence" (2009: 113). This section investigates three aspects of passive subjects as discourse-linkers in OE: information status, anaphoricity scores, and a comparison of passive subjects to active subjects. The data in this section are based on the same two texts that were used for the study in the previous section: *Orosius* and *Lives of Saints*.

Table 4.6 below repeats the data from Table 4.5 above, but now only includes the information status of the subjects.

Table 4.6 Information status of subjects in long passives in *Orosius* and *Lives of Saints*

Information status	<i>n</i>	%
OLD	56	88.9
ACCESSIBLE	0	0.0
ANCHORED	2	3.2
NEW	5	7.9
Total	63	100.0

The table shows that the overwhelming majority of subjects in passives in *Lives of Saints* and *Orosius* are OLD (88.9%), while only a small percentage are ANCHORED (3.2%) or NEW (7.9%). As a further investigation of the information status, Table 4.7 shows the anaphoricity scores, which indicate whether a subject has an antecedent in the preceding 5 main clauses, either in the immediately preceding clause ('1 main clause'), or in the second up to fifth preceding clause ('2-5 main clauses'), or whether it is not mentioned before ('No antecedent').

Table 4.7 Anaphoricity results for subjects of long passives in *Orosius* and *Lives of Saints*

Distance to antecedent	<i>n</i>	%
1 main clause	38	60.3
2-5 main clauses	11	17.5
No antecedent	15	23.8
Total	63	100.0

These scores show interesting differences with the previous table. While roughly three-quarters of the subjects (49 examples, 77.8%) have an antecedent in the preceding 5 clauses, which means they are clearly OLD, the category of non-anaphoric subjects is larger than the category of ANCHORED and NEW subjects in the previous table, with 23.8% against 7.9%. This reflects the difference between the two measures: information status does not take distance of the antecedent into account, while the anaphoricity scores aim to capture exactly that. What the table also shows is that if there is a link to the preceding discourse, it is likely to be 'local': over half (60.3%) of the referents have an antecedent in the immediately preceding main clause; the category of antecedent in earlier main clauses is much smaller (17.5%). The following examples illustrate.

- (41) 'A certain man fell on ice and broke his arm, and lay in bed very severely afflicted, until some one fetched to him, from the aforesaid cross, some part of the moss with which it was overgrown,'
and se adliga sona on slæpe wearð gehæled on ðære ylcan nihte þurh Oswoldes gearnungum.
 and the sick-man was forthwith in sleep healed in the same night, through Oswald's merits
 (coelive [Oswald]:34.5408)

- (42) =(18)
 '...and they led Abraham's brother's son Lot, with his family, out of the city; [...]'
and Loth se rihtwisa wearð ahred ðurh God.
 'and Lot, the righteous, was delivered by God'
 (coelive [Pr_Moses]:211.2980)

The subjects of these sentences, *se adliga* in (41) and *Loth se rihtwisa* in (42) are both discourse-old, but where the antecedent for *se adliga* is in the immediately preceding sentence, the antecedent for *Loth se rihtwisa* is just outside the window of 5 clauses. The analysis of the information status and anaphoricity scores for the passive subjects, then, clearly shows that the majority of these subjects are indeed both given and anaphoric.

A focus on the discourse-linking function of passives (cf. also Mair 1990, see Section 3.4.1.1) carries with it an assumption that there could be differences between these passive subjects and active subjects in terms of their information status or anaphoricity. The first implication of this assumption is the idea that passive subjects are discourse-old, according to the notion that the main motivation to use them is to put a given argument in subject position. As an approximation, Table 4.8 presents the relative frequency of pronominal and ellipted subjects—in all likelihood representing given information⁵³—in active main clauses, versus those in passive main clauses.⁵⁴ The selection of clauses does not take into account the position of the subject, but only considers the number of pronominal subjects with respect to nominal subjects.

Table 4.8 Percentage of pronominal and ellipted subjects in passive and active main clauses in *Orosius* and *Lives of Saints*

	Total no. of subjects	Pronom. subjects	% Pronom.	Ellipted subjects	% Ellipted
	<i>n</i>	<i>n</i>	%	<i>n</i>	%
Active clauses	8161	1714	21.0	2441	29.9
Passive clauses	708	151	21.3	76	10.7

Pronominal subjects active/passive: $\chi^2(1) = .04$, $p = .84$; Ellipted subjects active/passive: $\chi^2 = 117.86$, $p = 0$

The table shows that the percentage of pronominal subjects in passive and active main clauses is very similar: 21.0% in active clauses against 21.3% in passives. The percentage for ellipted subjects, however, shows a considerable difference: 29.9% of the active subjects are ellipted, against 10.7% of the passive subjects. These results, then, do provide some evidence that passive subjects are given more often than active subjects, because they are more often ellipted and ellipsis is only possible when the subject is given. At the same time, we need to take into account the general tendency for subjects to be given, irrespective of the voice of the clause or the semantic role of the subject.⁵⁵

⁵³ I also selected subjects consisting of only a D-element, but those numbers and percentages are very low and there is only a small difference between active and passive (1.7% and 2.3%, respectively).

⁵⁴ The following queries were used: **Active main clauses**: IP-MAT idoms subject (NP-NOM|NP-NOM-RSP|NP-NOM-1) and a finite verb (*VBD*|VBP*|*MDD*|*MDP*|*AXD*|*AXP*|*BED*|*BEP*|*HVP*|*HVD*); **Passive main clauses**: IP-MAT idoms subject, a finite form of *to be* (*BED*|*BEP*) and a past participle (*VBN*); **Pronominal subjects**: (on the output of one of the preceding queries): IP-MAT idoms subject and subject idomsonly *PRO*; **Ellipted subjects**: IP-MAT idoms subject and subject idomsonly *con*. For all queries, traces were not ignored.

⁵⁵ *Orosius* and *Lives of Saints* behave quite differently from other texts. The percentage of pronominal subjects in active clauses in the entire YCOE is much higher, at 31.0%, while the percentage of ellipted subjects in active clauses is lower, at 24.8%. The percentages for the passive clauses, interestingly, do not show those types of differences, with 23.6% pronominal subjects, and 6.3% ellipted.

If we approach this from the opposite direction, we would expect that there are few new subjects among passive subjects. As an approximation of new information, Table 4.9 presents the frequency of typically new subjects: those modified by *sum*, as in (43), a number as in (44), or a quantifier, as in (45).

- (43) *Sum wer wæs betogen þæt he wære on stole,*
 ‘A certain man was accused of stealing’
 (coelive [Swithun]:265.4385)
- (44) *Ða þa Seruius Fuluius & Flaccus Quintus wæron consulas, wearþ on Rome an cild geboren, þæt hæfde IIII fet & IIII handa, & IIII eagan, & IIII earan.*
 ‘When Servius Fulvius Flaccus, and Quintus Calpurnius Piso were consuls, a child was born in Rome, that had four feet, and four hands, and four eyes, and four ears.’
 (coorosiu,Or_5:2.116.23.2444)
- (45) *Fela halige menn fram frymðe middaneardes wæron beforan us wundorlice gebogene,*
 many holy men from the-beginning-of-the-world were, before us, wondrously perfected
 (coelive [Ash_Wed]:279.2857)

One type of passive clause was excluded in this study of new subjects, namely late subjects, i.e. those sentences in which the subject follows both the finite verb and the passive participle, because they are clearly a separate category (see Warner 2007 and the next section). When these clauses are excluded, the total number of remaining clauses is 8161 active and 630 passive clauses; these were used to calculate the percentages in Table 4.9.

Table 4.9 Typically new subjects in active and passive main clauses in *Orosius* and *Lives of Saints* (excluding late subjects)

	Active	%	Passive	%
Subject with <i>sum</i>	151	1.8	11	1.7
Subject with number	157	1.9	20	3.2
Subject with quantifier	443	5.4	58	9.2

Active vs passive, *sum*: $\chi^2(1) = .04$, $p = .85$; with number: $\chi^2(1) = 4.64$, $p < .05$; quantifier: $\chi^2(1) = 9.34$, $p < .005$

The table shows that, contrary to the prediction, new subjects are slightly more frequent among the passive than among the active subjects. The differences overall are not large—and only statistically significant for subjects with a number or a quantifier—and it is not possible to tell here whether there are specific explanations

for individual examples in the table, but this suggests that passive subjects are not more often OLD than active subjects. It needs to be noted, however, that the category of quantifiers may not be as reliable a clue for new subjects as the other two; it includes the *sum*-subjects and contains some examples with old subjects, as in (46).

- (46) *and hi ealle wurdon anes dægges gehælede þurh þæs halgan ðingunge,*
and they all were one day healed through the saint's intercession
'and they were all healed in one day through the saint's intercession'
(coelive [Swithun]:187.4336)

This sentence has a quantifier, *ealle*, as part of its subject, but the subject is, by virtue of the pronoun, clearly discourse-old. Still, the data in the table do not provide evidence that passive subjects are less often new than active subjects. Again, if we consider the general tendency for subjects to be given, whether active or passive, these results are not surprising. However, it remains to be seen whether any of these tendencies change over time.

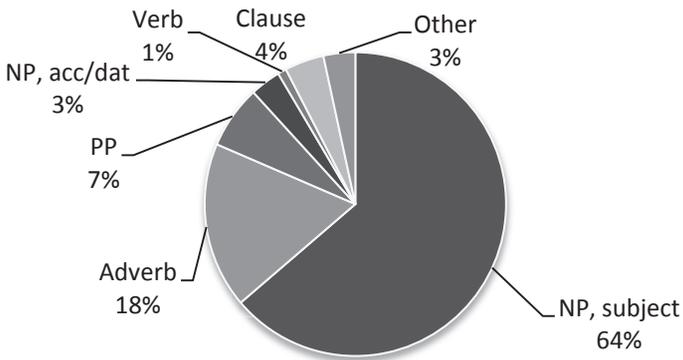
The second implication of the notion that passive and active subjects behave differently in terms of information structure because of the discourse-linking of passive subjects lies in the number of clause-initial subjects. If passive subjects are specifically used as unmarked themes from ME onwards, they should increasingly occur in clause-initial position, with the subject as the only theme, i.e. the given element. The question is, then, whether passive subjects occur less or more often than active subjects in clause-initial position, i.e. whether there is an indication of a difference in this indication for discourse-linking between active and passive subjects in OE. Table 4.10 shows the clause-initial elements in active and passive sentences in *Orosius* and *Lives of Saints*, in addition illustrated with the pie charts in Figures 4.1 and 4.2.

Table 4.10 Clause-initial elements in active and passive main clauses in *Orosius* and *Lives of Saints*⁵⁶

	Active	%	Passive	%
NP, subject	5212	63.8	371	52.8
Adverb	1445	17.7	170	24.2
PP	550	6.7	70	10.0
NP, no subject	268	3.3	6	0.9
Verb	76	0.9	21	3.0
Clause	344	4.2	21	3.0
Other	279	3.4	43	6.1
Total	8174	100.0	702	100.0

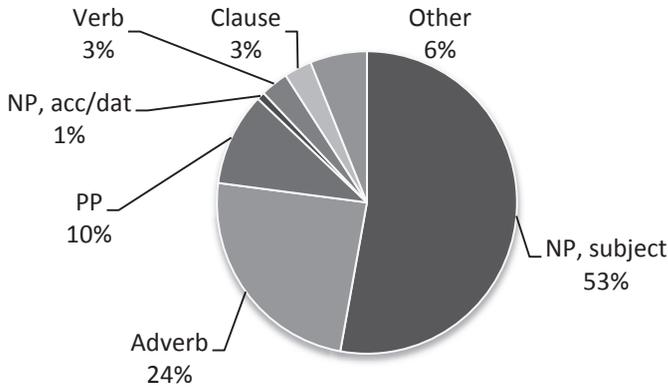
$\chi^2(6) = 90.05, p = 0$

Figure 4.1 Clause-initial elements in active main clauses



⁵⁶ The same queries for active and passive main clauses were used as before, now again including the post-VBN subjects. The following simple query was used in a series of queries, substituting the relevant element, to determine the clause-initial elements: IP-MAT idomsfirst subject/PP*/ADVP*/NP-ACC*/NP-DAT*/VB*/CP*/IP*. The complement file of these queries was used to determine the number of 'other' elements. Conjunctions were ignored. Note: as for NP-NOM, subjects are included in the category 'NP, subject', other types, such as left dislocations, in 'Other'.

Figure 4.2 Clause-initial elements in passive main clauses



The table and figures show that the passive clauses are in fact less often subject-initial than active clauses: 52.8% of the passive clauses are subject-initial against 63.8% of the active clauses. The elements that occur in clause-initial position in non-subject initial clauses are similar in passive and active clauses, and their order of frequency is mostly similar: adverbs are the second-largest category in both clause types, followed by PPs. The remaining categories are small, between 0.9% and 6.1%. These data suggest that, at least in OE, there is no clear indication that the passive subject is more often used as a discourse-linker than the active subject, because it occurs with other material, presumably mostly given, in the Left Periphery more often than active subjects.

In conclusion, this section has shown that in the OE data there is limited evidence to suggest that the information status of passive subjects is different from that of active subjects. While it is true that the subjects of the long passives are indeed mostly given and anaphoric—as a scenario about passives as discourse-linkers would suggest—the subjects of active and passive main clauses in fact show the same percentage of definitely given (pronouns) and presumably new (e.g. quantified) subjects, i.e. there are no indications that passive and active subjects behave differently in this respect. The only exception are the ellipsed subjects, whose percentage is considerably higher in active subjects than in passive subjects. Contrary to expectation, there are indications that passive subjects appear in clause-initial position less often than active subjects in OE, which suggests that in OE the passive subjects are not used more often than active subjects for local anchoring. As we have seen, the passive *by*-phrase can also function as a local anchor, which may explain the lower percentage of subjects in initial position. Section 6.2.2.1 follows up on this issue and presents the data for EModE to answer the question whether there are any changes in the comparative information status of passive and active subjects, or in the number of clause-initial subjects in passive compared to active clauses.

4.2.2.4 Late subjects

Recall from Section 2.2.1 that passives are one type of clause in which subjects are sometimes found to occur late in the clause (cf. Van Kemenade 1997, Warner 2007). The following example illustrates such a late subject.

- (47) *and in þis bataile was slayne **Nemion**, þat was Cassibalanus broþer*
 and in this battle was slain Nemion that was Cassibalanus' brother
 'and in this battle, Nemion, who was Cassibalanus' brother was killed'
 (Brut 32.14; Warner 2007: 92, 10b)

In (47) the subject occurs after both the finite verb *was* and the passive participle *slayne* and presumably stays within the VP. Passives thus provide a distinctive word order option in that they allow for a subject to surface in this late position in the clause, a property they share with unaccusative verbs. Passives, however, have the advantage that the position of the subject is easier to determine because of the diagnostic presence of the participle, while this is more difficult for unaccusatives with a single verb, as in (48).

- (48) *Of þese seuene heuedes comen **alle manere of synnes***
 from these seven heads come all manner of sins
 'From these seven heads spring all manner of sins'
 (Vices and Vertues 11.8; Warner 2007: 94, his 11)

As Warner (2007: 93) points out, it is not possible to tell whether the subject is in the immediately postverbal position (the V2 position, i.e. Spec,TP) or whether it has remained within the VP. Note that in some instances, it is in fact possible to determine the position of a subject of an unaccusative as late because of other VP material preceding the subject, as illustrated in (49).

- (49) *Than entered onto þe castell **on Jon Butler***
 then entered into the castle one John Butler
 'Then one John Butler entered into the castle'
 (Capgrave Chronicle 239.23; Warner 2007: 92, 10f)

The syntactic structure of these clauses presents an information-structural opportunity for speakers to place new information later in the clause: indeed, it has been proposed that late subjects are always (or overwhelmingly often) new—see Biberauer & Van Kemenade (2011) for OE and Warner (2007) for ME. Biberauer & Van Kemenade's (2011: 23) formal analysis makes a very clear prediction: all high passive subjects should be old, all late passive subjects should be new.

Let us turn to late subjects in passive clauses in the present corpus in order to establish whether they are indeed mostly new. The initial selection for late subjects

was based on Query 4.1, i.e. clauses with a form of *to be* and a past participle, to which the requirement was added that the clause should have an overt subject and that this subject should follow the past participle. Note that the subject should not just follow the finite verb, but specifically the participle because, under the assumption that it remains in V, it is a diagnostic for the position of the subject, because any subject that follows it is necessarily in its initial position in the VP. I analysed the information status, anaphoricity and persistence of these subjects, according to the categories outlined in section 4.2.2.1 above.

There are a total of 78 late subjects in the two texts: 64 examples in *Lives of Saints* (12.7% of the 502 passive clauses), and 14 in *Orosius* (6.2% of 226 passive clauses). One type of sentence was excluded from the selection because it represents a fixed pattern which is likely to have its own specific function, as illustrated in (50): these clauses start with an appositive noun, and the subject proper occurs after the past participle.

(50) *Anastasius wæs gehaten se arwurpa mæssepreost þe se bisceop to fundode swa færllice mid gange.*

Anastasius was called the venerable masspriest, whom the bishop to went so quickly in journey

‘Anastasius was the name of the venerable mass-priest, to whom the bishop went so quickly in his journey’

(coelive [Basil]:466.785)

This pattern accounts for almost a quarter of the examples (18 examples, 23.1%); there are only two verbs that occur in these sentences: *gehaten* (16) and *gecyged* (2). These sentences introduce a new referent in this late, post-participle position, and make it available for topic status in the following sentence and thus function as a type of topic introducing construction (cf. Lambrecht 1994: 176-177). Indeed, these sentences often occur in the opening of a saints’ life or a new paragraph within such a saints’ life, where they are clearly used to introduce a new protagonist in the story. In terms of the information status, all but three subjects are NEW, as in (50). The three sentences in which the subject is not NEW are presented in (51)-(53).

(51) *Mannases wæs gecyged Ezechias sunu,*

‘Hezekiah’s son was called Manasses,’

(coelive [Book_of_Kings]:434.3972)

(52) *Iohannes wæs gecyged þæs Symones sunu, se wæs æfter his fæder ðæs folces heretoga,*

‘John was chosen, the son of Simon, who was after his father, the people’s leader,’

(coelive [Maccabees]:741.5315)

- (53) *'A certain shoemaker* was sewing the holy man's shoes, and pierced through his hand very severely, but the holy man healed him immediately, and inclined him to baptism, and many others with him.'
- Anianus wæs gehaten se ylca sutere,*
 'This same shoemaker was called Anianus.'
 (coelive [Mark]:23.3207-10;27.3211)

The subjects in (51) and (52) are ANCHORED, meaning that even though they are not strictly speaking discourse-new, they would be classified as 'new' in a binary system and crucially, even though they are linked to the discourse, the referents themselves are new to the discourse. Example (53) is more puzzling. The subject is OLD, referring back to the referent in the preceding sentences, and it seems that the name, rather than the person, is introduced to the discourse, albeit in an unexpected position. As a result the name is focused, although it does not become clear from the context why emphasis or contrast would be appropriate here. The scores for anaphoricity and persistence make clear that the majority of these 18 sentences indeed represent cases of topic introduction: 15 of the 18 examples have a subject that is not mentioned in the preceding sentences, but which is subsequently picked up in the following sentence.

Table 4.11 presents the analysis of the information status of the remaining 60 subjects.

Table 4.11 Information status of late subjects in *Lives of Saints & Orosius*

Information status	<i>n</i>	%
OLD	16	26.7
ACCESSIBLE	10	16.7
ANCHORED	2	3.3
NEW	32	53.3
Total	60	100.0

The table shows that with 53.3%, more than half of the subjects are NEW. This is clearly a much higher percentage than the previous results for subjects—only 7.9% of the subjects of long passives were NEW (see Table 4.6). Moreover, it is quite a deviation from the general observations on the givenness of subjects (cf. Section 2.4.1). Late subjects, then, clearly behave differently from subjects in other positions and seem to represent a distinct option for structuring information in the sentence. One telling illustration of this principle are the *there*-clauses, which clearly introduce, or present, a new element late in the clause.

- (54) *And þær wurdon ofslagene eahta hund wera,*
 'and there were slain eight hundred men' (coelive [Maccabees]:320.5044)

- (55) *þær wurdon ofslagene sume þreo þusend,*
'There were slain some three thousand'
(coelive [Maccabees]:357.5066)
- (56) *and þær wurdon gehælede þurh his halgan geearnunge fela mettrume menn fram mislicum cōpum.*
'And there were healed through his holy merits many infirm men of various diseases'
(coelive [Oswald]:190.5491)
- (57) *þær wurdon gehælede þurh ðone halgan wer feower wanhale menn innan ðrym dagum,*
'Then there were healed, by the holy man, four sick men within three days'
(coelive [Swithun]:143.4303)

With respect to the predictions in the literature, however, it is also clear that this generalization does not accurately describe the information-structural function of all late subjects, because there is still a considerable percentage of subjects that are OLD (26.7%) or ACCESSIBLE (16.7%), suggesting that introducing or just expressing new referents is not the only function of these late subject constructions, and that there may be other, syntactic, factors that determine the position of the subject.

It is possible to identify several such factors in the database that influence the position of the subject, which may either go hand in hand with information-structural factors or overrule them. First, almost half of the OLD subjects (7 of 16) occur after *then*-group adverbs, known to show inversion almost invariably (see Chapter 2, Section 2.2.2), meaning that in these sentences the subject is postverbal for syntactic rather than information-structural reasons. Examples (58) and (59) illustrate.

- (58) *þa wearð gefullod se foresæda Nicostratus, mid his wife Zoe and þrym and ðrittigum mannum þe him ær folgodon, and mid him gefullode wæron.*
'There were baptized the aforesaid Nicostratus, with his wife Zoe, and three and (coelive [Sebastian]:127.1288)
- (59) *þonne bið oferswyðed swa eac seo galnyss.*
then will-be overcome also the adultery
(coelive [Memory_of_Saints]:325.3504)

Postverbal position in unaccusative clauses is of course not the same as late, as we have seen, but it means that in these cases the choice is not between the highest and one of the lower positions, but only between the two lower positions, Spec,TP and the late subject position. Another factor which is relevant is weight. Weight has been shown to often correlate with information status and is sometimes used as a

short-cut to information status (cf. Warner 2007: 82-83, referring to Wasow 2002). Indeed, many of the late subjects are quite long, for instance in example (58) above. Information status and weight do not always correlate in the database, however, because it also contains examples of subjects which are long but OLD, as in example (60).

- (60) *þus wearð geopenad se arwurðe mæssepreost, and þæt halige mæden þe his huses gymde.*

‘Thus was made manifest the venerable mass-priest, and the holy virgin who took care of his house;’

(coelive [Basil]:490.802)

It could be that weight is an independent factor influencing the position of the subject as late, or that in some cases it goes hand in hand with information status. A final factor that seems to be relevant to the OLD late subject is that six of the late subjects occur in verb-first clauses, as in the following examples.

- (61) *Wearð þa geopenad his earman wife his manfullan behat þam hetolan deofle.*

[there] was then revealed to-his poor wife his wicked promise to the hateful devil

(coelive [Basil]:405.726)

- (62) *And wæron gegeorcode þa reðan wyta.*

and [then] were made-ready the cruel tortures

(coelive [Eugenia]:196.311)

The reason for the subject occurring late in these examples could be due to the fact that these are verb-first sentences, which have a specific information-structural character, described as, for instance, presentational or introductory (see Hinterhölzl & Petrova 2010). The fact that the clauses are verb-initial may have a stronger influence on the position of the subject than the information status of the subject.

There is one further consideration, and that is the question whether new passive subjects always occur in this late position. In other words, is it a tendency that works both ways, in that new subjects tend to occur late and late subjects tend to be new? Table 4.9 already gave an indication that this is not straightforwardly so: it showed that some of the *sum*-subjects, which are generally new, do occur in one of the higher subject positions (in fact, 11 pre-participle against 15 post-participle). Table 4.6 also already showed that, even though it is not common, new subjects do in fact occur in the higher subject positions (7.9% of the non-late passive subjects were NEW).

The data from *Orosius* and *Lives of Saints* show that the late subject position is predominantly, but not exclusively, used by discourse-new subjects. There is a clear

preference for the late subjects to be new (or non-late subjects to be old—this is not clear), but there is also a considerable category of counterexamples, with roughly one-third of the late subjects in the database being OLD or ACCESSIBLE. The existence of the counterexamples means that while the tendency is clear, it is not possible to connect the late subject position exclusively with new information. This has implications for analyses such as Biberauer & Van Kemenade (2011), which formalize the tendency as a syntactic rule. In addition, certain syntactic factors also determine subject placement and some of these may obscure the effect of information status—such as verb-initial clauses—while others go hand in hand with information status—such as weight.

4.2.3 Summary and conclusions

This section has investigated the information-structural and syntactic properties of the passive in OE. The results show that the passive was already used as an information-rearranger in OE and that achieving given-before-new order was relevant to the passives in the selected texts. At the same time, it became clear both from the syntactic properties of the OE passive as well as the lower frequency of long passives that perhaps the passive is less important as an information-rearranging device than it was to become in later periods.

The section started with a discussion of the literature on the passive in OE, in order to select criteria for selecting the passives in the corpus study that followed. The passive in OE was shown to be a less standardized construction than it is in PDE, most clearly visible in the availability of different auxiliaries as well as a range of prepositions that could encode the agent phrase. Following the earlier conclusions in Chapter 2, I argued that the functional or semantic distinctions—between stative and dynamic, and adjectival, verbal and resultative passive—were to be disregarded in the corpus study, because the information-structural aspects are relevant across the different types.

Three aspects of passives in *Orosius* and *Lives of Saints* were examined. The first corpus study aimed to establish the percentage of passives that followed given-before-new order and showed that the vast majority (roughly 90%) of long passives obeyed the given-before-new principle. The second study focused on passive subjects as discourse linkers and found that both active and passive subjects are important as discourse linkers—there were no indications that this is true for passive subjects to a larger extent than for active subjects. The third study focused on late subjects, which showed a preference for new subjects, but it also became clear that other factors (like fixed constructions and weight) played an important role in determining the position of the subject as late or early in the sentence.

It is clear that the main information-structuring function of long passives in OE is already that of creating a given-before-new order. It shares with short passives the discourse-linking properties of the subject, where it fits into the existing properties for subjects in general. In terms of subject positions, the passive is a particularly useful construction in that it structurally allows for a subject to remain low, which in

turn can be used for information-structural purposes, namely to place a new subject in a position later in the sentence.

Section 4.3 Object fronting

This section investigates object fronting in OE and focuses on two aspects: first, the information-rearranging function of object fronting, which makes it a possible alternative to the passive in OE; and second, the information-structural characteristics of the fronted object with respect to the discourse-linking function of the first position and those of the subject with respect to the interaction between information status and inversion. In this section, I show that (i) object fronting is used as an information-rearranging device, showing given-before new order of information of object and subject; (ii) fronted objects in clause-initial position have an important discourse-linking function but also a separate function of hosting contrastive elements; and (iii) although the information status of the subject cannot be directly linked to a position in the sentence, subject placement follows a general given-before-new pattern.

Object fronting, a syntactic operation known most commonly as (object) topicalization, is closely associated with focus or contrast in PDE, and is usually information-structurally marked. Speyer (2010) distinguishes the following three types of object fronting for PDE.⁵⁷

- (63) Beans he likes, but peas he hates.
- (64) Pterodactylus, it is called.
- (65) This proposal, we discussed at length.

(Speyer 2010: 30, his 4b,d,c)

Example (63) shows what Speyer calls double-focus-topicalization and (64) shows focus movement. Unlike the first two types, the fronted element in (65), an example of ‘anaphoric preposing’, is not contrastive and it presents discourse-old information. The fronted element is what Speyer calls a “propositional topic”, i.e. it refers to a proposition, “the subject matter” of a couple of sentences in the preceding discourse, rather than a specific referent (2010: 30).

The character of object fronting in OE is crucially different from that in PDE, witness the additional type of topicalization that Speyer (2010) identifies specifically for OE. He characterizes it as the fronting of an “aboutness e[ntity]-topic” (2010: 38), so in contrast to anaphoric preposing, this topic does have a specific referent. The aspect that stands out is that the fronted element contains topical information (in

⁵⁷ Speyer also includes one type of topicalization of a non-object, which he calls scene-setting preposing (as in *In the afternoon, I usually go for a walk*): fronting of a prepositional or adverbial phrase. This type of topicalization will not be included in this section because of the focus on object fronting.

the sense of aboutness topic) and it is not contrastive, unlike the topicalizations in (63) and (64). Examples (66) and (67) illustrate this type of object fronting.

(66) *þone asende se Sunu,*

this sent the Son

‘The son sent this one’

(coaelhom,ÆHom_9:114.1350; Speyer 2010: 38, his 17a)

(67) *ne hine ne drehð nan ðing,*

and-not him not troubled no thing

‘and nothing troubled him’

(coaelhom,ÆHom_11:558.1780; Speyer 2010: 38, his 17b)

Even though this is a type of ordering that is defined by Speyer in terms of aboutness topic rather than givenness, considering the close relation between topichood and givenness, it carries implications for the given-before-new principle. Siewierska (1984: 218) defines this same type of object fronting as enabling the most topical (in the sense of sentence-topic, see Lambrecht 1994: 131) and/or given (discourse-old) argument to occur in clause-initial position.⁵⁸ She points to the existence of this strategy in German and also points out that English lacks such strategies, with the exception of the passive.

It is in this respect—the fact that object fronting can place an old element in clause-initial position and crucially before the subject—that objects are relevant as alternatives to the information-rearranging function of passives. Consider the following two examples, two consecutive entries from *Orosius*, both describing the reign of a king and stating that he was killed by his own men. But where the first example has a long passive, the second has a fronted object.

(68) *Æfter þæm þe Romeburg getimbred wæs DCCCC wintra & LXII, feng his sunu to rice Antonius, and hit [hæfde] VII gear. He hæfde twa gesweostor him to wifum. He hæfde folc gegaderad, & wolde winnan on Parthe, ac he wearð ofslagen on þæm færehte from his agnum monnum.*

‘Nine hundred and sixty-two years after the building of Rome, his son, Antoninus, succeeded to the empire, and held it [not full] seven years. He had two sisters for his wives. He had gathered an army, and wished to fight against the Parthians; but, in the march, he was put to death by his own men.’
(coorosiu,Or_6:17.142.17.2993)

⁵⁸ Siewierska calls this “topicalization”, but I will not use this term as it is confusing between the syntactic operation of (object) fronting—placing an element in initial position—and a specific information-structural function of placing topical information (regardless of the syntactic status) in initial position in the clause.

(69) *Æfter þæm þe Romeburg getimbred wæs DCCCC wintra & LXX, feng Marcus Aurelius to Romana onwalde, & hiene hæfde IIII ger. Hiene ofslogon eac his agene men, & his modor mid.*

‘Nine hundred and seventy years after the building of Rome, Marcus Aurelius succeeded to the government of the Romans, and held it four years. His own men slew him, and also his mother.’

(coorosiu, Or_6:17.142.22.2996)

These examples have a similar context and message since they both introduce an emperor and report how he was killed by his own men, in the same logical structure: ‘the emperor’ is the patient, and ‘his own men’ as agent. But while example (68) uses a passive to place the more given ‘emperor’ in first position, example (69) fronts the object to achieve the same effect. These examples suggest that, at least in some contexts, passivization and object fronting were indeed alternatives in OE.

There is an additional factor which is relevant to the use of object fronting, which is what has already been described in Chapter 2 as part of the given-before-new principle: the important function of old information in clause-initial position to link the sentence to the preceding discourse. The idea is that the first position is used specifically for discourse-linking (cf. Birner & Ward (1998), who point to the relevance of anaphoricity even in focus preposing and topicalization). This also ties in with Los’s (2009) claim that the first position in OE is especially used for unmarked themes, including but not limited to subjects, and that this is what makes it different from the use of the first position in PDE, where only subjects can be unmarked themes (cf. Halliday 1967). Note that all the OE fronted objects in examples (66), (67) and (69) contain discourse-old information, providing a link to the immediately preceding sentence.

This section is structured as follows. Section 4.4.1 discusses the selection of the sentences with fronted objects that were used for this study, illustrating the syntax of the relevant sentences and explaining why some fronted objects are not relevant to the present study. Section 4.4.2 presents the results of the study of the relative information status of the fronted objects and subjects in the database, and aims to establish whether object fronting was used as an information-rearranging device, and to what extent there is an overlap with the passive. Section 4.4.3 focuses in more detail on the fronted objects, and investigates their relevance as discourse linkers in clause-initial position. Section 4.4.4 focuses on the subjects in the object fronting clauses and investigates the information-structural properties of pre- and postverbal subjects, in an attempt to determine whether there is a relation between the position of the subject and its information status. Section 4.4.5 concludes.

4.3.1 Method and overall frequencies

This section describes the selection of fronted objects for the corpus study and the method of analysis. Query 4.3 below was used to select the fronted objects in

Orosius and Lives of Saints. It selects all main clauses with a subject and object in which the object is the clause-initial element.

Query 4.3 Fronted objects
a. IP-MAT idoms subject (defined as NP-NOM NP-NOM-RSP NP-NOM-1)
b. IP-MAT idoms object (defined as NP-ACC NP-ACC-RSP NP-ACC-RFL NP-ACC-1 NP-ACC-RSP-1 NP-ACC-RFL-1)
c. IP-MAT idomsfirst object

The query states that the clause should have both a subject (line 4.3a) and an object (line 4.3b) as main clause constituents, with the object in clause-initial position (line 4.3c). Conjunct clauses were excluded or included where necessary by adding the node *CONJ* to the command 'add_to_ignore' in *CorpusStudio*; when this label is part of the 'add_to_ignore' command, conjunct clauses are included because the programme does not 'count' them as clause-initial elements in query line 4.3c.

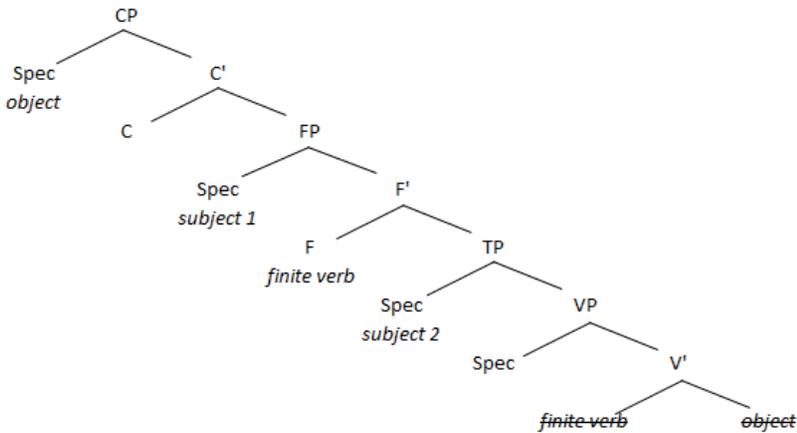
Examples (70) and (71) below illustrate the two orders that this query selects, with (70) showing inversion of subject and finite verb and (71) showing the uninverted order. In this and the following examples, the fronted objects are given in bold, and the subjects, where relevant, are underlined.

- (70) *Twægen þissera dæla habbað deor and nytenu mid us, þæt is gewylnunge and yrre.*
 two of-these parts have beasts and cattle [in common] with us, that is [to say], desire and anger;
 (coelive [Christmas]:100.79)

- (71) *and **ðas feower godspelleras** God geswutelode gefyrn, on ðære ealdan æ Ezechihele þam witegan.*
 and these four Evangelists God revealed before, in the Old Law, to-Ezekiel the prophet
 (coelive [Mark]:174.3311)

Figure 4.2 illustrates the analysis of these two word order options, following the general analysis of OE main clause syntax as outlined in Chapter 2.

Figure 4.2 Inversion and non-inversion after clause-initial objects



Because the clause-initial object in Spec,CP is not one of the operator(-like) clause-initial elements that invariably trigger inversion, the finite verb does not move up to C, but only to F. This also means that these sentences show the variable type of V2: the subject occurs either before the finite verb, in Spec,FP, or after the verb, in Spec,TP, giving rise to V2 and V3 sentences, respectively.

4.3.1.1 Other object-subject orders

There are other clauses in which the object precedes the subject and finite verb,⁵⁹ but these were excluded from the corpus study for various reasons. The first type includes clauses which show the same order as the clauses in the preceding section, but have an additional XP preceding the object. Depending on the type of this XP, the finite verb occurs either in C, preceding the object, or in F, following the subject. Example (72) shows one of the 26 XVOS orders in the database with the operator-like *þa* in clause-initial position, triggering inversion; example (73) shows one of the 16 cases of XOVS order, with a PP in clause-initial position and the verb in a lower position, after the object.

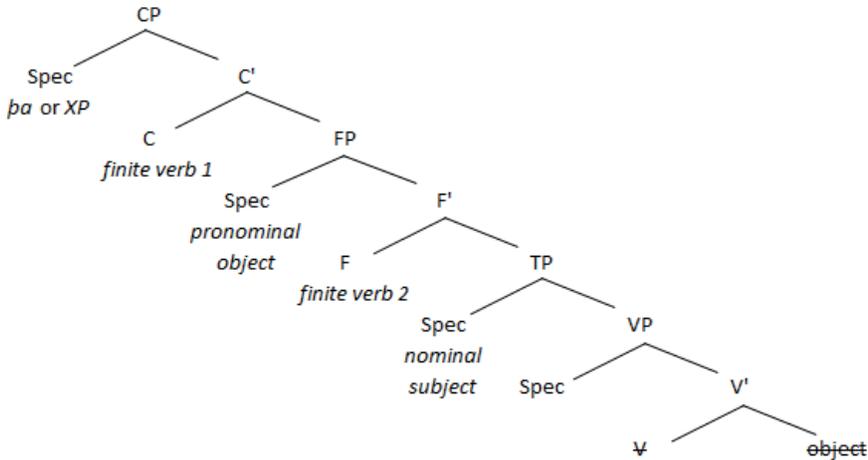
- (72) *þa gefeng hine an treow be ðam fexe sona, forðan þe he wæs sidfæxedede*
then caught him a tree by the hair immediately, because he was long-haired
(coelive [Alban]:218.4125)

⁵⁹ These were selected using Query 4.3, lines a and b, with an additional query that the object should precede the subject, but without line 4.3c stipulating that the object should be clause-initial.

(73) & on *ðære hwile þe he þær winnende wæs, frefelice hiene gesohte*
ſMinotheo, seo Scibbisce cwen, mid þrim hunde wifmonna,
 and on the time that he there fighting was, boldly him sought Minothæa the
 Scythian queen, with three hundred women
 ‘and, while he was fighting there, Minothæa, the Scythian queen, with three
 hundred women, boldly sought him out’
 (coorosiu,Or_3:9.71.4.1391)

In these examples, the object is pronominal and occurs in Spec,FP, the position where pronominal subjects normally occur (Van Kemenade 1987, Fischer et al. 2000: 118-120). In the case of an operator in Spec,CP (24 times *þa* in the examples with this order in the database, cf. (72), once *ne* and once *þær*), the verb is in C, and the surface order is XVOS. If the XP in Spec,CP is a non-operator, as in (73), the verb moves to F and the surface order is XOVS. Figure 4.3 illustrates the two different orders.

Figure 4.3 X-initial clauses with fronted objects



Both these orders have been excluded from the study in this chapter because the objects are not in clause-initial position and the focus of this section lies specifically with those clause-initial objects in order to gain insight into the use of the clause-initial position. In the case of the XVOS examples, there is another issue in view of the study in Section 4.3.4, which is that inversion is triggered by the presence of an operator *þa*, i.e. inversion is necessarily grammatically rather than information-structurally driven, which makes the information status of the subject irrelevant.

There is one other order that occurs in the database, which is less uniform than the patterns above; this is surface XOSV, which occurs 16 times in the two texts. It comes in two types. The first type, with 8 examples, has a pronominal object, *man* as subject, and an XP in first position.

- (74) *for ðy hit man hæet Wislemuða.*
 therefore it they call mouth-of-the-Vistula
 ‘therefore, they call it the mouth of the Vistula’
 (coorosiu,Or_1:1.16.32.315)
- (75) *& þeh hit mon hæet eal Syria.*
 and though it they call all Syria
 ‘and though it is all called Syria’
 (coorosiu,Or_1:1.10.16.141)
- (76) *For þon hi mon hæet on Crecisc Amazonas, þæt is on Englisc fortende.*
 therefore them they call in Greek Amazons, that is in English, seared
 ‘they were, therefore, called in Greek Amazons, that is in English, seared’
 (coorosiu,Or_1:10.29.34.582)

Note that these types of object fronting, with *man* as the subject, represent a clear alternative to the passive in terms of their impersonalization function, and that (75) and (76) are in fact translated as passives in Bosworth (1859). Los (2002: 192-193) points out that OE *man* “exhibits the same functional equivalence with passives as Dutch *men*” and provides two examples of Latin passive sentences which are translated into English with a *man* sentence, even though, as Los points out, a passive would be possible in OE.

- (77) Lat.: Eos qui ducuntur ad mortem eruere ne cesses
 Those who led^{PASS3PL} to death free^{INF} not hesitate^{2SG}

OE translation:

þa þe man læt to deaðe alys hi ut symble
 those whom people lead to death free them out always
 ‘always set free those who are being led to death’
 (Æls (Edmund) 214; Los 2002: 192, her 19)

- (78) Lat.: Sicut Filius hominis non venit ministrari sed ministrare
 so son of-man not came serve^{PASSINF} but serve^{ACTINF}

OE translation:

Swa mannes sunu ne com þæt him man þenode ac þæt he þenode
 so of-man son not came that him people served but that he served
 ‘So did the Son of Man not come to be served but to serve’
 (Mt (WSCp) 20.28; Los 2002: 193, her 20)

Even though these orders represent a very clear alternative to the passive, they were excluded for the present study because the sentences are not object-initial, and *man*

subjects are not relevant to the information-structural analysis, as they are always ACCESSIBLE but cannot be anaphoric.

The other 8 XOSV examples present a more mixed collection of sentences. The subjects are pronominal or proper nouns (twice), and the clause-initial X can be any type, an adverb as in (79), a subordinate clause as in (80), or a both an adverb and a PP as in (81).

- (79) *and eac ure fynd he het us lufian,*
and also our enemies he commanded us to-love
'and likewise He commanded us to love our enemies'
(coaelive [Memory_of_Saints]:259.3468)
- (80) *ealle Asiam hie genieddon þæt hie him gafol guldon,*
all Asia they forced that they them taxes paid
'They forced all Asia to pay them taxes'
(coorosiu,Or_1:10.29.7.563)
- (81) *& þa æt nehstan hyre agene sunu hio genam hyre to geligere;*
and then, at last, her own son she took her to lie-with
'Then, at last, she took her own son to her bed'
(coorosiu,Or_1:2.22.24.445)

One interesting observation is that the objects occur in a high position, where normally pronominal objects occur, because Spec,CP is presumably occupied by the adverb or PP (4 examples are in fact pronominal, as expected). This raises questions for the syntactic analysis of these sentences, for instance whether the clause-initial elements are in fact extra-clausal or whether there is an additional position in the Left Periphery. However, from the perspective of the present study, the additional adverbial in front of the object makes these sentences irrelevant, for the same reason that the other XP-initial sentences were excluded.

4.3.1.2 Frequency of object fronting in the database

Table 4.12 presents the frequencies of fronted objects and the other subject-object orders against all main clauses with a subject and an object in *Orosius* and *Lives of Saints* (selected with lines a and b of Query 4.3 above and with both conjunct and non-conjunct clauses included). The table also contains the percentages of the same queries for all the texts included in the YCOE and the subperiods that the two selected texts belong to, O2 (850-950) and O3 (950-1050).

Table 4.12 Fronted objects and other object-subject orders in texts from the YCOE

	Active transitive	Fronted objects	% Fronted objects	Other Obj-Sbj orders	% Other Obj-Sbj orders
	<i>n</i>	<i>n</i>	%	<i>n</i>	%
<i>Orosius</i>	612	80	13.1	34	5.6
O2	2004	258	12.9	142	7.1
<i>Lives of Saints</i>	929	75	8.1	24	2.6
O3	7151	660	9.2	240	3.4
Entire YCOE	13573	1403	10.3	582	4.3

Between *Orosius* and *Lives of Saints*, fronted objects: $\chi^2 (1) = 10.19, p = .005$

The table shows that object fronting is slightly more frequent in the O2-text *Orosius* than in the O3-text *Lives of Saints*, with 13.1% against 8.1% for the object-initial clauses. The other types of clauses in which the object precedes the subject are also more frequent in *Orosius* than in *Lives of Saints*, 5.6% compared to 2.6%. The percentages for the two texts are comparable to the percentages for all the texts in their sub-period (13.1% against 12.9% for *Orosius*; 8.1% against 9.2% for *Lives of Saints*), which means that the difference between the two texts reflect a more general difference between the two subperiods, with a decrease in the percentage of fronted objects from O2 to O3. The percentages for the fronted objects are also reasonably similar to those given in Speyer (2010: 27), who gives an average of 11.2% for the entire YCOE: 11.9% for OE1-2 (which includes *Orosius*) and 10.8% for OE3-4 (which includes *Lives of Saints*).⁶⁰

4.3.1.3 Information status and anaphoricity categories, including ‘speech’, ‘event’ and ‘topic’

The method for the information status and anaphoricity categories is similar to the method in the study of the passive in Section 4.2. Table 4.13 repeats the information status categories.

Table 4.13 (=Table 4.2) Annotation scheme for information status

Category	Categories from other models included in each category	Binary distinction
OLD	Prince’s (1981) <i>Textually Evoked</i> , Komen et al.’s (2014) <i>Identity</i> , Pintzuk & Taylor’s (2011, 2014) <i>Previously mentioned</i> .	‘Given’

⁶⁰ Speyer (2010) does not give his queries, which means that it is not possible to determine in detail whether exactly the same clauses were included.

ACCESSIBLE	Prince's <i>Unused</i> and <i>Situationally Evoked</i> , Komen et al.'s <i>Assumed</i> , Taylor & Pintzuk's <i>Shared/Cultural knowledge</i> , <i>Situationally Evoked</i> , Haug et al.'s (2014) <i>Generic</i> , Birner's (2006) <i>Elaborating inferables</i>	'Given'
ANCHORED	Birner's <i>Bridging inferables</i> , Prince's <i>Brand-new Anchored</i>	'New'
NEW	Prince's <i>Brand-New Unanchored</i> , Taylor & Pintzuk's and Haug et al.'s <i>Specific, Non-specific</i> and <i>Short-term referents</i> , Komen et al.'s <i>Inert</i>	

The measure of anaphoricity—previously consisting only of ‘no antecedent’ and ‘prior mention’ with the addition of the distance to the antecedent in number of clauses—is expanded with ‘general’ anaphoricity references. These are necessary for those cases where there is clearly a reference to the previous discourse, but it is not possible to find one specific antecedent in the preceding text to which the object refers. I divided these general references into three categories, illustrated in (82)–(84).

- (82) ‘Elisha also healed a nobleman from the dreadful leprosy; he was called Naaman, of the land of Syria, and he believed in God through the great miracle which God wrought in him.’

Fela oðre tacna gefremede God þurh hine
many other signs performed God through him
(coelive [Book_of_Kings]:313.3892)

- (83) They said that Cornelius sent them to him; ‘He is a Centurion, and hath the fear of God, a very righteous man,’

þæt wat eall þeos scyr.
that knows all this province
‘... as all this province knows;’
(coelive [Peter's_Chair]:119.2357)

- (84) ‘One natural law is appointed to all mankind, that no man may do harm to another man, even as the Saviour said in His holy gospel; ‘That which thou desirest not to befall thyself in thy life, that do not to another man.’

þis cwæð Drihten sylf.
‘This said the Lord Himself.’
(coelive [Forty_Soldiers]:348.2699)

In (82), the preceding sentence ends the description of a particular miracle, although the word for ‘sign’ or ‘miracle’ is never mentioned. It is clear from the form of *Fela oðre tacna* that it refers back, but it can only refer back to the entire event; these references were labeled ‘event’. In (83) the reference is equally general: it refers to

the topic under discussion, and technically refers back to the entire sentence; these were labeled 'topic'. The reference is also general in (84), but this time it refers to a quotation in the previous sentence(s); these were labeled 'speech'. In all cases the distance for anaphoricity is analysed as '1': the last mention of, or last piece of the event, topic or speech is in the immediately preceding clause, even though they may have been continued for more clauses before that.

4.3.2 Relative information status of fronted object and subject

This section aims to establish to what extent the object-initial clauses in the two selected texts follow the given-before-new order of information and thus to what extent they function as information-rearrangers. Table 4.14 presents the results of the analysis of the information status of the fronted object and the subject, including both inverted and uninverted sentences, and conjunct as well as non-conjunct clauses. Three sentences with negative subjects and/or objects were excluded because of their problematic information status. As in the table for long passives in Section 4.2.2.2, the rows show the information status of the first element, the object in this case, while the columns show the information status of the second element, the subject in this case. The percentages with respect to the total number of examples are given between brackets and the shaded cells indicate those instances where the object and subject have an equal information status, or the object (the first element) is more given.

Table 4.14 Relative information status of fronted object and subject in object-initial clauses

		Subject				Total
		OLD	ACCESSIBLE	ANCHORED	NEW	
Object	OLD	42.8%(65)	20.4% (31)	3.9% (6)	4.6% (7)	71.7% (109)
	ACCESSIBLE	8.6% (13)	2.0% (3)			10.5% (16)
	ANCHORED	5.3% (8)		2.0% (3)		7.2% (11)
	NEW	9.9% (15)	0.7% (1)			10.5% (16)
Total		66.4 (101)	23.0% (35)	5.9% (9)	4.6% (7)	100.0%(152)

$\chi^2 (9) = 22.31, p < .01$

The table shows that 75.7% (115 out of 152) of the examples occur in the shaded cells, meaning that the fronted object and the subject have the same information status, or the object is more given, following a given-before-new order. These examples represent a clear majority, but the percentage is lower than that for the passive clauses (88.9%), suggesting that maintaining given-before-new order is not the only function of object fronting. Another difference with the data for the long passives is that the subject, as the second element, is in general more given (66.4% of the subjects in Table 4.14 are OLD, against 42.9% of the agent phrases in Table

4.5), which is not surprising when we consider the general tendency of subjects to be given or at least accessible. Of the 115 examples in the grey cells, 71 (46.7% of all examples) have an identical information status for the fronted object and the subject, for example OLD/OLD as in (85) or ACCESSIBLE/ACCESSIBLE as in (86).

- (85) ‘The soul is the mistress of the body, and governs the five senses of the body, as out of a royal throne. These senses are thus named: *Visus*, that is, Sight; *Auditus*, Hearing; *Gustus*, Taste with the mouth; *Odoratus*, Smelling with the nose; *Tactus*, Touching or feeling with all the limbs, but most usually with the hands.’

Ðas fif andgitu gewisseð seo sawul to hire wyllan,

these five senses directs the soul to its will

‘The Soul directed these five senses according to its will’

(coaelive [Christmas]:202.161)

- (86) ‘The island, Sicily, is three-cornered. At each corner there are hills: the north corner is called Pelorus, near to which is the city Messina: the south corner is called Pachynum, near which is the city Syracuse;’

& *þone westsceatan man hæt Libeum;*

and the west-corner they call Lilybæum

‘and the west corner is called Lilybæum’

(coorosiu,Or_1:1.21.5.416)

In (85), both *Ðas fif andgitu* and *seo sawul* are mentioned in the preceding sentences. In (86), *þone westsceatan* is inferable from the announcement of ‘at each corner’ in the previous sentence and the mention of the north and south corner, and *man* is generic, so they are both ACCESSIBLE.

The rest of the 115 examples which fall into the shaded cells (44 examples, 28.9% of the examples in total) are examples of a fronted object which is more given than the subject, and so these sentences really show a given-before-new order of information. The following examples illustrate the orders OLD/ACCESSIBLE, OLD/ANCHORED and OLD/NEW, respectively.

- (87) ‘He is a Centurion, and hath the fear of God, a very righteous man,’

þæt wat eall þeos scyr.

‘as all this province knows;’

(coaelive [Peter's_Chair]:119.2357)

- (88) ‘They would then have soon perished before their enemies, if they had not broken into the city by a device, which was most shameful, though it was afterwards thought most worthy of them; [...].’

***Pysne nyttan cræft**, þeh he arlic nære, funde heora tictator, Camillis hatte.*
this useful device, though it honourable not-was, found-out their dictator,
Camillus called

‘This useful device, though it was not honourable, was found out by their Dictator, Camillus.’

(coorosiu,Or_2:8.51.24.983)

- (89) ‘Lo then! Quintianus, Christ’s adversary, went in a ship over Semithetus (the river Symæthus) about Agatha’s possessions, desiring also to apprehend all her kindred, but he could not for Christ.’

***Hine gelæhte an hors**, þa ða he læg on ðam scipe, hetelice mid toðum*

‘A horse seized him, as he lay in the ship, savagely with its teeth’

(coaelive [Agatha]:211.2145)

In all three examples, the fronted object is OLD: a demonstrative in (87), a NP containing a demonstrative and lexical repetition in (13), and a pronoun in (89). The subjects belong to different categories: *eall þeos scyr* in (87) is ACCESSIBLE—it is not mentioned before, but inferable from the situation; *heora tictator, Camillis hatte* in (13) is ANCHORED—it has both a possessive pronoun to anchor it and the addition of the name shows that it is not assumed to be accessible by the writer; *an hors*, the subject of (89), is NEW—it has an indefinite article and is not mentioned in the preceding discourse, there is no anchor, and again, the explanation in the relative clause following the subject indicates that it is not assumed to be accessible information.

The table also shows that the remaining 37 of the 152 (24.3%) examples have an unexpected order of information from the perspective of the given-before-new principle, with the fronted object being less given than the subject. A considerable number of these examples (15 examples) are clearly contrastive, in the sense that alternatives are present in the text, see Krifka (2007). In some of these contrastive examples, as in (90) and (91), the two alternatives occurring after one another show the same object fronting and a similar violation of the given-before-new principle.

- (90) ‘Then went he through all the Egyptian land, sowing God’s seed, and healed the sick.

***Hreoflige he** geclænsode fram ðære unclænan coðe, **Wode he** gehælde*

the lepers he cleansed from the unclean disease; the possessed he healed

(coaelive [Mark]:6.3189)

- (91) ‘Then the bishop was greatly astonished, and took the housel which the Saviour had blessed, brake [it] in three, and consumed one portion;’
***Pone oðerne dæl** he dyde gehealden mid him to bebyrgenne æfter his forðsiðe. **Pone ðryddan dæl** he dyde onsondor,*
 the second portion he caused to-be-kept with to be buried after his death;
 the third portion he caused to-be-set-apart
 (coelive [Basil]:123.531 and 125.532)

In (90), *Hreoflige* is contrasted with *the possessed* in the following sentence, while in 0, *Pone oðerne dæl* is contrasted with *one portion* in the preceding sentence. The large number of contrastive examples in the selection of given-before-new exceptions suggests that a deviation from the given-before-new order takes place when the writer wants to focus something, i.e. when there is the alternative motivation for use of the first position (cf. Lambrecht 1994: 31-32 on the competition between unmarked topics and marked foci in the clause-initial position). What is important here is the definition of contrast in terms of alternatives that are present in the text, even though not always as literally or explicitly as in examples (90) and 0. Example (92) shows an ACCESSIBLE/OLD order of information and (93) ANCHORED/OLD.

- (92) ‘Elisha also healed a nobleman from the dreadful leprosy; he was called Naaman, of the land of Syria, and he believed in God through the great miracle which God wrought in him.’
***Fela oðre tacna** gefremede God þurh hine*
 many other signs performed God through him
 (coelive [Book_of_Kings]:313.3892)
- (93) ‘Six hundred and sixty years before the building of Rome there was that very great battle between the Cretans and the Athenians. The Cretans gained the bloody battle,’
***& ealla þa æðelestan bearn para Atheniensa** hi genoman,*
 and all the most.noble children of.the Athenians they took
 ‘and they took all the most noble children of the Athenians.’
 (coorosiu,Or_1:9.28.14.547)

In (92), *Fela oðre tacna* is ANCHORED and it is contrasted with the event described in the previous sentence (‘the great miracle’), while in (93), the Cretans take the ‘most noble’ children, but not the other children. In this last example, the contrast is inferable rather than explicitly mentioned in the text.

Not all of the 37 violations of the given-before-new order are clearly contrastive in the sense that alternatives are present in or inferable from the text. Two categories can be distinguished among the remaining examples. The first category of examples seems to show a different type of focus, or of marked

information structure, namely emphasis. Consider the following examples, all showing NEW/OLD order of information.

- (94) *Swa micela mihte he hæfde menn to gehælenne,*
 so great might he had men to heal
 ‘So great might had he to heal men’
 (coelive [Martin]:484.6270)
- (95) ‘However, we will tell you of a certain thane, who lived thirty years with his wife in continence;’
Pry suna he gestrynde,
 three sons he begat
 (coelive [Æthelthryth]:125.4213)
- (96) *Twa mila hæfde Martinus fram his mynstre to Turonian byrig þær se bisceopstol wæs,*
 two miles has Martin from his monastery to Tours city where his episcopal-seat was
 ‘Martin had two miles to go from his monastery to the city of Tours in which was his episcopal seat;’
 (coelive [Martin]:1198.6764)

The first thing to note in these examples is that all but (95) open a new section or paragraph. In (94), the addition of *swa micela* indicates emphasis lexically and means that it carries more emphasis than if it had just been *mihte* (note also that in the PDE translation, the quantification *so great* is a trigger for I-to-C movement, cf. Section 2.2.2). In (95), there seems to be a contrast or focus on the three sons because of the context of a story about abstinence: he *first* had three sons but *then* they lived in abstinence. In (96), although there is no clear indication for whether *Twa mila* needs to be considered as a great or small distance, either way it could be focused: ‘it was only two miles so he went there often’ or ‘it was as big a distance as two miles but still this great man went there as often as he could’.

The second category of examples that do not follow a given-before-new order of information and are not contrastive raises questions with regard to the nature of the distinctions between certain information status categories. While examples such as (94)-(96) show a clear violation of the given-before-new principle with a NEW object occurring before an OLD subject, other objects have an information status which is only one category less given than the subject. Example 0 and (98) illustrate an ACCESSIBLE/OLD order of information.

- (97) 'It was so much wasted by that burning, that it never afterwards was such [as it had been], till Augustus, in the year when Christ was born, rebuilt it so much better, than it ever was before, that some men said, it was adorned with precious stones.'

þone fultum & þæt weorc *Agustus* *gebohte mid fela M talentana.*

that help and that work Augustus paid-for with many thousand talents
(coorosiu,Or_6:1.133.19.2816)

- (98) *þa halgan apostolas Petrum and Paulum* *he* *geseah* *gelome,*
the holy apostles Peter and Paul he saw frequently
(coaelive [Martin]:688.6414)

The fact that there is no clear reason why there should be a violation of the given-before-new order in these sentences could be an indication that *þone fultum & þæt weorc* and *þa halgan apostolas Petrum and Paulum* should be considered as OLD rather than ACCESSIBLE, in which case there would be no violation. *þone fultum & þæt weorc* were coded as ACCESSIBLE because there is no specific or clearly inferable antecedent; the description of the building process as 'help' is newly added and is not present in the earlier sentences. *Peter and Paul*, on the other hand, were analysed as ACCESSIBLE because they do not occur in the text but are part of the speaker's and hearer's world knowledge. As another type of explanation, it could also be the case that the unexpected order is allowed because the categories OLD and ACCESSIBLE are not, in fact, hierarchically ordered, in the sense that ACCESSIBLE referents (or perhaps only a subset of ACCESSIBLE referents) cannot be considered to be less given than the OLD subjects. Remember that the category ACCESSIBLE contains both referents that are expected to be accessible because of world knowledge or because of inference from the discourse. It could be that there is a difference and world knowledge should belong to the category OLD. It is not possible to answer the question on the basis of these data (but cf. Taylor & Pintzuk 2014, who find a small difference between the different types of what belongs to the current ACCESSIBLE category, although they also find that accessible referents behave differently from discourse-old referents).

It becomes clear from this study that the clear majority (75.7%) of object fronting sentences show an order of information of the fronted object and the subject that follows a given-before-new pattern, suggesting that object fronting is indeed used in OE to rearrange arguments, and thus information, according to a given-before-new principle. In that sense, then, object fronting truly is an alternative to the passive for this information-structural function. What makes object fronting different from the passive, however, is that there is a distinct second function for object fronting, with the majority of the 24.3% exceptions to the given-before-new order being information-structurally marked. This markedness shows itself in the contrastive or emphatic nature of the fronted object, similar to the functions of object topicalization described for PDE. This study, then, illustrates Lambrecht's

(1994: 31-32) observations about the first position as showing competition between unmarked themes and marked foci.

4.3.3 The first position as a local anchoring position

Having established that object fronting indeed has a function of rearranging information, this section focuses in more detail on the clause-initial objects, and aims to establish their information-structural function in the important first position in the clause. The study takes into account not only the information status of the clause-initial object, which was already partly considered in the previous section, but also its anaphoricity and persistence. The data in this section show that an important function of the clause-initial objects is that of ‘local anchoring’: providing an unmarked (i.e. non-contrastive) link to the immediately preceding sentence.

First of all, let us again consider the information status of the fronted objects, but now without the comparison to the information status of the subject. Table 4.15 repeats the results for the fronted objects from Table 4.14.

Table 4.15 Information status of fronted objects

Information status	<i>n</i>	%
OLD	108	71.5
ACCESSIBLE	16	10.6
ANCHORED	11	7.3
NEW	16	10.6
Total	151	100.0

The table shows that the majority of objects are OLD (71.5%), but that there are quite a number of examples that belong to the two ‘new’ categories, with 7.3% ANCHORED and 10.6% NEW examples. It already became clear from the analysis of the examples that did not follow a given-before-new order of information in the previous section that a number of these ANCHORED or NEW PPs in clause-initial position are contrastive or otherwise focused. It is not the case, however, that there is a direct relation between newness and contrast/focus: new objects can also be non-contrastive, and likewise, given objects can be contrastive.

A second issue is the anaphoricity of the clause-initial object. Table 4.16 presents the results of the analysis of the distance between the fronted object and its antecedent, as well as the type of reference. Note that the category ‘no antecedent’ in the table below includes not only the NEW objects, but also OLD referents which are not included in the five preceding clauses (but may have been mentioned before those five clauses), which is why the category of ‘no antecedent’ is larger than the 27 NEW or ANCHORED objects in the table above.

Table 4.16 Anaphoricity results for clause-initial objects in main clauses

Distance to antecedent	<i>n</i>	%
1 clause (event, speech, topic)	23	15.2
1 clause (object, subject, anchor)	83	55.0
2-5 clauses	14	9.3
No antecedent	31	20.5
Total	151	100.0

Table 4.16 shows that the majority of fronted objects have an antecedent in the immediately preceding clauses: the top three rows make up 79.5% of all the examples. The majority of antecedents for these examples are in the preceding clause, while the percentage of antecedents for which the distance is larger than 1 is small, only 9.3%. These results show that when there is a link to the preceding discourse, this linking takes place very locally, providing evidence that ‘local anchoring’ is an important function of the clause-initial object. If we compare these results to the anaphoricity results for the subjects in the same clauses, the difference in linking behaviour between the first position and positions later in the clause becomes even clearer.

Table 4.17 Anaphoricity results for subjects in object-initial main clauses

Distance to antecedent	<i>n</i>	%
1 clause (event, speech, topic)	0	0.0
1 clause (object, subject, anchor)	54	35.8
2-5 clauses	26	17.2
No antecedent	71	47.0
Total	151	100.0

The percentage of subjects referring to an antecedent in the immediately preceding clause is considerably smaller: 35.8% compared to 70.2% for the fronted objects. What is also interesting is the much higher percentage for a longer distance to the antecedent (17.2% against 9.3% for objects) and for ‘no antecedent’ (47.0% against 20.5%). Not only do these results show that the subjects in these clauses behave differently from the fronted objects with respect to anaphoricity, the results are especially interesting because the majority of subjects in these sentences are in fact OLD, as we have seen in the previous section (66.4%, Table 4.14, with an additional 23.0% being ACCESSIBLE). This means that these subjects, showing a somewhat similar percentage of OLD referents to the objects (both around 70%), act differently specifically with respect to anaphoricity, in turn highlighting the importance of the (local) linking or anchoring function of the clause-initial objects.

If we finally look at the scores for persistence for the fronted objects, we see that, similar to anaphoricity, persistence is either local or absent: the middle category is rare.

Table 4.18 Persistence results for clause-initial objects in main clauses

Distance to subsequent mention	<i>n</i>	%
1 clause	66	43.7
2-5 clauses	5	3.3
No subsequent mention	80	53.0
Total	151	100.0

The ‘no subsequent mention’ category accounts for just over half of the examples. If we compare this to the tables above, this seems to suggest that in many cases linking is indeed important for the objects in fronted position, much more so than keeping a referent in the discourse. This is especially striking for the ‘event’, ‘speech’ and ‘topic’ categories: none of these 15 examples have a subsequent mention. Roughly one-third of all examples (35.6%) follow a similar pattern: they have a referent in the immediately preceding clause but no subsequent mention. Again, these results point to the importance of local anchoring in the first position: this is where the link to the previous discourse is made, and while linking back seems a common pattern, keeping the referent as the topic of the discourse happens less often.

In terms of combined results for anaphoricity and persistence, there is another pattern that stands out—these are the cases where the fronted object is the object in the preceding and the following clause (5 examples), or where object is the subject of the preceding and the following clause (2 examples). It seems that in both cases fronting of an object is used as a way to maintain the same characters in the same grammatical position or, alternatively, in the same role, i.e. the same aboutness-topic. Consider the following two examples.

- (99) ‘In their power, the Samnites were so bold, that the prince called Pontius, who was their leader, told them to ask the king, his father, who was at home, whether he would rather that he should kill them all, or order them while living to be put to shame.’

Hie þa se æþeling to ðæm bismre getawade þe þa on ðæm dagum mæst wæs, þæt he hie bereafade heora claþa & heora wæpna, & siex hund gisla on his geweald underfeng, on þæt gerad þæt hie him sibban ece þeowas wæren.
 ‘The prince then tortured them with the shame, which was the greatest in those days,—he stripped them of their clothes, and their weapons; and took six hundred hostages into his power, with the view, that afterwards, they should always be his slaves. (coorosiu,Or_3:8.66.25.1299)

- (100) ‘He made all things and continueth from everlasting to everlasting’
Hine ne mihte nan þing gewyrcean, for ðon þe nan þing næs ær he,
‘nothing could make Him, because nothing was before Him,’
(coaelive [Christmas]:66.53)

In (99), the fronted object *Hie* refers to *the Samnites*, who are the object in the preceding sentence, and are again picked up as the object in the following sentence. The point of this order seems to be to maintain the syntactic structure—*the Samnites* as the object and *prince Pontius* as the subject. In (100), it is the information structure that is kept the same in the three sentences: the fronted object *Hine*, referring to Christ, is the subject in the preceding and the following sentence, but in all three sentences it occurs as the clause-initial element. A fronted object, then, can either be used to maintain the syntactic structure or the order of elements, while in both cases an information structure of given-before-new is also achieved. Interestingly, example (100) is similar to an example given by Sato (2012: 639) to show Ælfric’s highly planned style: he uses the stylistic device chiasmus (B-A-A-B) to emphasize the elements that are different in the two clauses, the verbs.

To conclude, the main function of the objects in clause-initial position seems to be that of ‘local anchoring’. This is a function that makes it stand out from the characterizations of object topicalization in PDE. The notion of local anchoring is supported by the various results discussed in this section: the majority of fronted objects are discourse-old and non-contrastive, they provide a link to the immediately preceding sentence and while linking back is common, keeping the referent in the discourse after the fronted object clause seems to be less important. The comparison to the subjects occurring after the object, which are also mostly OLD or ACCESSIBLE, has provided further proof that anaphoricity is especially important for the clause-initial position: these subjects are less often anaphoric and the distance to the antecedent is generally larger. In addition to the function of local anchoring, the results also further confirm the findings from the previous section that a considerable number of the fronted objects (roughly a quarter) are similar to PDE object topicalization in that they are contrastive and/or present new information. It is clear, however, that these sentences represent a minority pattern in OE.

4.3.4 Inversion in object fronting sentences

Where the previous section focused on the fronted objects, this section focuses on the subjects following those fronted objects. Specifically, it investigates the predictions presented in the literature (e.g. Van Kemenade & Westergaard 2012, Hinterhölzl & Petrova 2010, see Chapter 2) that subject placement in OE—specifically, the variation between the two subject positions in Spec,FP and Spec,TP, resulting in either V2 or V3 surface word order—is determined by the information status of the subject. In this section, I show that given subjects tend to occur preverbally and new subjects tend to occur postverbally, confirming these predictions, but that there are a considerable number of counterexamples.

Table 4.19 presents the overall percentage of inversion in the two texts. For this study, object fronting clauses with nominal subjects were selected, excluding both pronominal subjects and *man*, because *man* generally behaves like a pronoun (Van Bergen 2000). I have also made a distinction between conjunct and non-conjunct clauses because these have been shown to behave differently, with conjunct clauses showing less inversion (see e.g. Bech 2001, but see Bech (in prep)). Note that the structural ambiguity between two types of postverbal subjects—inverted subjects in Spec,TP and late subjects which have stayed in the VP (cf. sections 4.2.2.4 and 4.4.3)—does not play a role for fronted object clauses because they are transitive and the grammatical subject is the external argument in all cases (i.e. originating in Spec,vP and not as an original internal argument within the VP).

Table 4.19 Inversion with nominal subjects in object-initial main clauses

		Inversion	Total	% Inversion
Conjunct clauses	O2	9	36	25.0
	O3	56	86	65.1
	YCOE	110	201	54.7
Non-conjunct clauses	O2	30	66	45.5
	O3	125	164	76.2
	YCOE	212	328	64.6

Entire YCOE, between conjunct and non-conjunct clauses: $\chi^2(1) = 5.14, p < .05$

The table shows that there is a difference between the two subperiods, with O2 below the average for the YCOE and O3 above the average for the entire YCOE, for both conjunct and non-conjunct clauses: O2 has 25.0% and O3 65.1% for conjunct clauses against an average of 54.7% for the entire YCOE, and O2 has 45.5% and O3 76.2% for non-conjunct clauses against an average of 64.6% for the entire YCOE. These data show that inversion is less frequent in conjunct clauses than in non-conjunct clauses, confirming earlier observations in the literature (but see Bech in prep.).

Having established the rate of inversion in the entire YCOE, and specifically the two subperiods to which *Lives of Saints* and *Orosius* belong, we can now turn to the rates of inversion in those two texts. Table 4.20 presents the results.

Table 4.20 Inversion with nominal subjects in object-initial main clauses in *Lives of Saints* and *Orosius*

		Inversion	Total	% Inversion
<i>Orosius</i>	Non-conjunct	8	13	61.5
	Conjunct	3	22	13.6
	Total	11	35	31.4
<i>Lives of Saints</i>	Non-conjunct	16	17	94.1
	Conjunct	13	15	86.7
	Total	29	32	90.6
Total of two texts	Non-conjunct and conjunct	40	67	59.7

Total numbers, between the two texts: $\chi^2 (1) = 24.35, p < .001$

The table provides a number of details about inversion in these two specific texts. First, it is clear that the rates of inversion differ quite dramatically in both texts: 31.4% in *Orosius*, against 90.6% in *Lives of Saints*. Second, the table shows that the rates of inversion in conjunct clauses differs from that in non-conjunct clauses and while in *Lives of Saints*, the conjunct clauses still show a similar pattern of high inversion (86.7%), in *Orosius* the rate of inversion becomes much lower in conjunct clauses. Because the numbers are small, I will not distinguish between conjunct clauses and non-conjunct clauses when considering the information-structural aspects of inversion and I will also not distinguish between the two texts, even though they might also show variation in information-structural motivation for subject placement.

Table 4.21 shows the distribution of the four information status categories of the subject in inversion and non-inversion sentences.

Table 4.21 Information status of nominal subjects in object-initial main clauses in inversion and non-inversion contexts in *Lives of Saints* and *Orosius*

	Inversion		Non-inversion		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
OLD	22	52.4	20	47.6	42	100.0
ACCESSIBLE	4	44.4	5	55.6	9	100.0
Subtotal 'Given'	26	51.0	25	49.0	51	100.0
ANCHORED	8	88.9	1	11.1	9	100.0
NEW	5	71.4	2	28.7	7	100.0
Subtotal 'New'	13	81.3	3	18.7	16	100.0
Total	39	58.2	28	41.8	67	100.0

For all four categories: $\chi^2 (3) = 5.27, p = .15$; for the binary distinction given/new: $\chi^2 (1) = 4.59, p < .05$

The table shows quite a clear tendency for NEW subjects to be inverted, i.e. to occur after the finite verb (71.4%), but it does not show the reversed pattern for OLD subjects: only 47.6% show uninverted order. If we only consider the binary distinction between ‘given’ and ‘new’, the tendency for ‘new’ subjects to be inverted becomes even clearer, with the greater part (81.3%) of these subjects occurring after the finite verb. The tendency for the ‘given’ subjects does not become more evident, with 51.0% inversion. In other words, ‘new’ subjects are likely to occur after the finite verb, but ‘old’ subjects are just as likely to occur before the finite verb as after the finite verb.

Let us consider some of the exceptions in these results to the general expectation that preverbal subjects are old and postverbal subjects new. First, examples (101)-(103) present the two NEW and one ANCHORED subjects that are not inverted.

(101) *Ac hiene Gandes seo Œea þæs oferfæreldes longe gelette, for þæm þe þær scipa næron*

but them Gyndes the river going-over long hindered, because there ships not-were.

‘But the river Gyndes long hindered him from going over, because there were not any boats there.’

(coorosiu,Or_2:4.43.5.809)

(102) *ac hi Creacas þær onfundon,*

but them Greeks there found

‘but there the Greeks found them’

(coorosiu,Or_1:4.23.15.461)

(103) *& hiene his agen ealdormon Artabatus besirede & ofslog.*

and him his own chief-officer Artabanus plotted-against and slew

‘and his chief officer Artabanus plotted against him, and slew him.’

(coorosiu,Or_2:5.48.24.934)

The subjects in (101) and (102) were analysed as NEW because the referents had not been mentioned in the text before. The subject in (102), *Creacas*, presents an interesting case because in this case it seems that the Greeks should be considered as at least accessible precisely because they occur in the position where most given subjects occur, while if they had occurred after the finite verb, it could have been a case of topic introduction, in which case they would have seemed newer (although newness is no requirement for topic introduction, cf. Lambrecht (1994: 176-177) on topic promotion). In (103), the fact that *his agen ealdormon Artabatus* occurs before the finite verb again could indicate that this is not a bridging inferable but an elaborating inferable, i.e. every commander has chief officers and so it can behave as a ‘given’ subject. As indicated for example (102), regardless of the information

status, the occurrence of these subjects before or after the subject may indicate their function in the sentence: as part of the background, or as a new topic. In the case of topic introduction, and perhaps also in the case of backgrounding information, this may not always be a question of information status.

What is perhaps more striking than the three inverted NEW or ANCHORED subjects is the large number of OLD (22) or ACCESSIBLE (4) subjects that occur in postverbal position. At least a part of these can be explained with reference to factors shown to influence inversion in OE or ME in earlier studies (e.g. Van Kemenade 1997, Warner 2007). One of these is the verb *say*, as in example (104), which accounts for 4 examples.

- (104) *bis cwæð Drihten sylf.*
'this said the Lord himself'
(coelive [Forty_Soldiers]:348.2699)

There are two further things that stand out in these inverted OLD subjects. The first is the high number (19) of examples with pronominal objects, as in (104) and the following example.

- (105) *Hine wylcumed se casere,*
him welcomed the emperor
'The emperor welcomed him'
(coelive [Agnes]:339.1946)

One way to explain this is from a perspective of language use: the finite verb functions as a divider of the two arguments, making it easier for hearer to identify them. Another way to explain this is in structural terms: if the pronominal object in these clauses is in Spec,FP rather than Spec,CP, i.e. the clause lacks a CP projection, the only position that is left for the subject is Spec,TP, the position after the finite verb. The second observation about these OLD inverted subjects is that a considerable number (13 of 26) are proper names. The information status of proper names is sometimes difficult to establish and often the fact that the name is used without any additional information seems to indicate that it is assumed to be accessible. On the other hand, they may need to be reactivated and for that reason occur later in the sentence, i.e. in postverbal position.

Whatever the explanation for these old inverted subjects, there is one final option that needs to be considered. Most of these exceptions (21 of 26) come from *Lives of Saints*. We have already seen that Ælfric is known for highly stylized writing and that *Lives of Saints* is very rhythmical, which may explain some of the exceptional patterns. More importantly, Table 4.20 showed that inversion is a clear majority pattern in *Lives of Saints* and it is likely that a syntactic pattern which is so strong, in a context where no other factors seem to influence the inversion (speakers are in fact able to distinguish between different contexts, cf. operator-

initial clauses), leaves less room for information-structural motivations. In other words, the syntax overrules information structure.

In conclusion, in this small database there are a considerable number of counterexamples to the proposal that there is a division between preverbal vs postverbal subjects in terms of information status. The database is only small but the number of exceptions suggests that in these two texts at least, information status seems to be only a secondary factor, amidst syntactic patterns and possibly other information-structural functions such as topic introduction. The tendency that can be found is that *NEW* and *ANCHORED* subjects are more likely to occur after the finite verb, while for *OLD* subjects there was no clear tendency to occur before the verb. This points both to a general tendency for old information to occur early in the sentence and/or new information to occur later in the sentence, and to the possibility that syntactic factors are, possibly specifically at this stage of the language, more important than information-structural factors, with inversion being the most frequent pattern in all clauses but the conjunct clauses in *Orosius*.

4.3.5 Summary and conclusions

This section aimed to establish to what extent object fronting was used as an information-rearranging device in OE. Additionally, it aimed to establish to what extent the first position was used for unmarked themes and to what extent subject placement was determined by information-structural factors. The results of the studies show that object fronting is indeed an alternative to the passive as an information-rearranger; that objects in clause-initial position for the majority function as local anchors; and that subject placement follows a general given-before-new pattern. The studies also showed that for all these general findings, there are exceptions, for instance in the clear alternative function of fronted objects of providing contrast.

The properties of object fronting in OE are crucially different from PDE in that object fronting is in most cases unmarked, i.e. non-contrastive. The data showed that the majority of object fronting clauses follow the given-before-new principle (75.7%), suggesting an overlap in function between passive and object fronting. Indeed, there are several examples of object fronting sentences with *man* as a subject which are clear alternatives, for more than information-structural reasons, to the passive. At the same time, fronted objects have a distinct separate function, witness the number of exceptions in the selection. These exceptions could be characterized as Lambrecht's marked foci: the objects are in the majority contrastive or emphatically focused. This second information-structural function of object fronting shows that the functional overlap with the passive is only partial: there is a clear alternative motivation for the use of fronted objects.

The second study showed that the function of the majority of clause-initial objects can be described as that of 'local anchoring': these objects provide an unmarked link to the immediately preceding sentence. This type of function of the fronted objects is a type that is only a minority pattern in PDE, if it exists at all. In

contrast, the majority pattern for PDE, contrastive or focused objects in clause-initial position, was only present as a minority pattern in OE. Considering the diachronic questions, the fact that this pattern was already in use in OE is significant because this shows that it was not a later introduction.

The final study of this section addressed the question of influence of information status on subject placement in the object-initial clauses. The study showed that there is a general tendency for old subjects to occur before the finite verb, while new subjects are more likely to occur in postverbal position. However, there were also a number of counterexamples to this tendency, which I explained with reference to a number of factors, such as the high percentage of inversion in the texts, meaning that information-structural factors are perhaps less important.

Section 4.4 Clause-initial PPs

This section investigates the information-structural properties of PP-initial clauses in the two selected OE texts. The aims are three-fold: (i) to establish to what extent clause-initial PPs represent local anchors; (ii) to examine the details of inversion after clause-initial PPs; and (iii) to establish to what extent variation in inversion after PPs is driven by information-structural factors. In this section, I show that the majority of clause-initial PPs in OE function as local anchors, but that they have additional functions of contrast and frame-setting. With respect to inversion, I show that the type of verb and the type of PP influence inversion, while information status seems to be a less important factor for subject placement in PP-initial clauses than it was in the object-initial sentences.

Clause-initial PPs in OE are of special interest because they represent the prototypical non-operator element described in the literature as showing inversion to a variable degree, giving rise to surface V2 and V3 orders. Moreover, in light of the two types of V2, it is relevant to note that PPs come in many different shapes and sizes—not only do they have a semantic versatility in that they can express different types of adverbials (such as time, place, manner, means), but they also have a syntactic versatility in that they can be both arguments and adjuncts. This is decidedly different from the more uniform clause-initial categories of objects and adverbs. It has not yet been established in the literature to what extent different types of PPs trigger inversion, except for a more general suggestion that arguments may trigger inversion more often than adverbials (see e.g. Warner 2007, Haeberli 2002: 251).

Clause-initial PPs represent one of the few contexts in which inversion can still occur in PDE, as illustrated in (106), although this use is restricted.

- (106) There are huge cartons and tins of nuts, vanilla, honey, peanut butter. Varieties of herbs tea are visible. *On the counter are loaves – whole wheat, cinnamon raisin, oatmeal, rye, soy sunflower, corn meal.* [Terkel 1974:607]
(Birner & Ward 1998: 156, their 198a)

This type of inversion has often been referred to as locative inversion, although not all initial PPs that show inversion are in fact locative, and not all locative PPs trigger inversion. In addition, this type of inversion is defined with reference to the types of verb with which it occurs: intransitive verbs, especially unaccusative verbs (see e.g. Levin & Rappaport 1995). Birner & Ward (1998) propose that this type of inversion functions as an information-rearranger, while it has also been proposed that locative inversion has a presentative function (Bolinger 1977, Bresnan 1994). Syntactically, the explanation for this occurrence of unaccusative subjects after the finite verb has been ascribed to their properties of originating as the internal argument within the VP and the possibility to remain in that position, like passive subjects (cf. Section 4.2.2.4).

Unlike objects, PPs are still quite frequent in clause-initial position in PDE, both with inversion and non-inversion, and although they, like objects, are mainly used contrastively, some functions seem to be slightly less marked than for clause-initial objects. For instance, Virtanen (1992) shows how the initial position, specifically for adverbials (including PPs) is used both for emphasis or contrast as well as textual cohesion, which she described as a more neutral function. At the same time, however, Biber et al. (1999: 803) show how only 5% of place adverbials occur in initial position, and consider presubject adverbials “very marked” (see also Hasselgård 2010). Krifka (2007), following Chafe (1976), describes frame-setting as an important function of clause-initial elements: providing a general context in which a statement should be interpreted. In Krifka’s definition, alternatives are always implied, which again suggests a function of the initial position as something resembling contrast, i.e. a marked use.

This section is structured as follows. Section 4.4.1 discusses the selection of clauses for the corpus study, showing how the annotation of the YCOE allows for a differentiation of types of PPs. Section 4.4.2 investigates the information-structural properties of the clause-initial PPs in an attempt to establish whether the function of clause-initial PPs, like clause-initial objects, is ‘local anchoring’. Section 4.4.3 investigates the rates of inversion after PPs in the entire YCOE, and in *Orosius* and *Lives of Saints* specifically, and aims to establish whether there are certain factors in terms of the type and/or form of the PP that determine whether inversion occurs. Section 4.4.4 asks the same question that was asked for subjects in object-initial clauses, whether the information status of the main clause subject correlates with the position in which it occurs. Section 4.4.5 concludes.

4.4.1 Method and selection

The way the PPs are annotated in the corpus means that more distinctions can be made automatically, which allows for selecting exactly those clauses that are relevant. Query 4.4 was used to select all main clauses that have a prepositional phrase as first constituent.

<i>Query 4.4</i> Clause-initial PPs
a. IP-MAT idoms subject and finite verb b. IP-MAT idomsfirst PP* c. PP* iprecedes finite verb or subject d. PP* idoms PP* CP* IP* (to be excluded from the results)
<u>Definitions</u> Finite verb: *VBD* VBP* *MDD* *MDP* *AXD* *AXP* *BED* *BEP* *HVP* *HVD* Subject: NP-NOM NP-NOM-RSP NP-NOM-#

Line 4.4a stipulates that the main clauses should have a subject and a finite verb, while line 4.4b selects only those clauses that have a PP in clause-initial position. Line 4.4c selects the clauses in which the PP is followed (without any intervening material) by either the finite verb (in inversion contexts) or the subject (in non-inversion contexts). This line was added to exclude sentences such as (107), in which the clause-initial PP is followed by another element (an ADVP in this example).

- (107) [_{PP} *On ðæm dagum þe Cirus Persa cyning Babylonia abræc*], [_{ADVP} *ða*] *wæs Croesus se liþa cyning mid firde gefaren Babylonium to fultume*;
at the time when Cyrus, Persians king, Babylon stormed, then was Croesus, the Lydians king, with army come Babylonians to help
‘At the time, when Cyrus, king of the Persians, stormed Babylon, Croesus, king of the Lydians, came with an army to help the Babylonians’
(coorosiu,Or_2:4.44.7.830)

The reason why these sentences were excluded is that it is likely to be the second element that triggers inversion, and not the initial PP. Another type of clause that was excluded is given in (108); in these clauses the initial PP contains two coordinated PPs.

- (108) [_{PP} *Ymb feower hunde wintra*] & [_{PP} *ymb feowertig þæs þe Troia Creca burg awested wæs*], *wearð Romeburg getimbred from twam gebroðrum, Remuse & Romuluse*.
about four hundred winters and about forty that when Troy Greek city laid-waste was, was Rome-city built by two brothers, Remus and Romulus
‘The city Rome was built by two brothers Remus and Romulus, about four hundred and forty years after Troy a city of the Greeks was laid waste.’
(coorosiu,Or_2:2.38.31.742)

These sentences were excluded because the two PPs are embedded within another PP, which complicates the execution of subsequent queries on the structure of the clause-initial PP. They were excluded by executing Query 4.4d on the results file of 4.4c and then using its complement file—all clauses which are not selected by

4.4d—for further queries. Line 4.4d also aims to exclude an idiosyncrasy of the YCOE annotation: some adverbial clauses in the YCOE are coded as PPs with a head P which contains the complementizer and an IP or CP complement, as in (109).

- (109) [IP-MAT *and* [PP [P *swa* [CP-CMP *hi hi gearnlicor sceawodon*]]], *swa scimodon heora eagon swiðor.*
 and so her they more-closely looked, so were-dazzled their eyes more
 ‘and the more closely they looked at her, the more their eyes were dazzled.’
 (coaelive [Agnes]:148.1813)

Finally, similar to the object fronting clauses, conjunct clauses were excluded or included where necessary by adding the node *CONJ* to the command “add_to_ignore” in *CorpusStudio*.

One of the areas of interest with respect to PPs is the form of the NP complement of the P, which can be distinguished in the corpus. Most importantly, it is possible to select the anaphoric elements inside these NP complements: demonstratives and pronouns. Query 4.5 presents the queries that were used to select NPs containing pronouns and demonstratives, either as the only element in the NP, or occurring together with other material.

<p><i>Query 4.5</i> The anaphoric elements in clause-initial PPs</p> <p>On clause-initial PPs from previous query:</p> <p>a. PP* idoms NP*</p> <p>b. NP* idomsonly D* OR NP* idomsonly *PRO*</p> <p>On complement file of the query above:</p> <p>c. PP* idoms NP*</p> <p>d. NP* idoms D* *PRO*</p>

Lines 4.5a and b select NP complements of the P that consist of only a demonstrative, as in (110), or a pronoun, as in (111).

- (110) *Mid þam wunode an mæden mærllice drohtnigende*
 ‘With him [lit. that] dwelt a virgin, living virtuously’
 (coaelive [Basil]:468.786)

- (111) *Æfter hiere feng to ðæm rice Penthesilia,*
 ‘After her Penthesileia took the sovereignty,’
 (coorosiu,Or_1:10.30.22.599)

Lines 4.5c and d select NP complements of the P that contain a demonstrative or pronoun and additional material, illustrated in (112) with an NP containing a

pronoun and a noun, and in (113) with an NP containing a demonstrative and a noun.

- (112) *Æfter ðysum wordum se eadiga Petrus gehadode Apollonarem and gehalgode to bisceope,*
 after these words the blessed Peter ordained Apollinaris and consecrated [him] as bishop
 (coelive [Apollinaris]:25.4540-1)

- (113) *Æfter his deaðe foran eft Cartainienses an Sicilie mid scipum.*
 after his death went again the Carthaginians to Sicily with ships
 (coorosiu,Or_4:5.91.29.1854)

Query line 4.5a was repeated on the complement file of lines 4.5c and d, to select all NP complements that do not contain a demonstrative or pronoun, such as (114).

- (114) *On Easterdagum he wolde etan fisc gif he hæfde.*
 ‘On Easter-days he would eat fish if he had [it].’
 (coelive [Martin]:1267.6800)

The clause-initial PP in (114) only consists of the preposition *On* and the bare noun *Easterdagum*.

Information status and anaphoricity categories, including ‘speech’, ‘event’ and ‘topic’

The fronted objects were analysed according to the same information status categories and categories for anaphoricity and persistence as outlined in Section 4.2.2.1. Like the fronted objects, clause-initial PPs often provide a general reference to the preceding discourse: ‘speech’, ‘event’ and ‘topic’ (cf. section 4.3.1.3). The following examples illustrate these categories for initial PPs.

- (115) ‘... and on them fell the curse which the prophet spake, [...]’
Be swilcum cwæð se hælend eac on sumere stowe,
 ‘Of such spoke the Saviour also in a certain place’
 (coelive [Mark]:119.3282)

- (116) ‘Thomas said to her: [...]’
Æfter þysum com hire wer to þam kyninge Migdeum
 ‘After this came her husband to the king Mazdai’
 (coelive [Thomas]:300.7725)

- (117) ‘Towards his fellow-soldiers, he had kindly feeling, and great love, and modest patience, and true humility above man’s measure. He had as great temperance in his food as if he had been a monk rather than a soldier;’

and for his æðelum þeawum his efencempan ealle þa hine arwurðodon mid wundorlicre lufe.

and for his noble qualities his fellow-soldiers all him revered with a marvellous love

(coaelive [Martin]:47.5997)

In (115), the reference is labelled ‘speech’: *swilcum* refers to a quotation spanning a couple of sentences preceding this PP-initial clause. In (116), the PP-initial clause is also preceded by a quotation of a couple of sentences, but in this example the reference is not to the content of the speech, but to the act of speaking. *bysum*, then, refers back to the events described in the previous clauses and is labelled ‘event’. In (117), finally, *æðelum þeawum* refers back to the discussion in the text just before this sentence, where these ‘noble qualities’ of Martin, the protagonist of this Saint’s Life, are described; this illustrates a ‘topic’ reference.

4.4.2 Clause-initial PPs as local anchors

The analysis of the information status and anaphoricity of the clause-initial objects in Section 4.3.3 showed that the primary function of clause-initial objects can be described as local anchoring: these objects contain a link to the immediately preceding discourse and they are information-structurally unmarked (i.e. not contrastive and not focused). This section aims to establish whether clause-initial PPs have the same function. The approach to the study of the information-structural character clause-initial PPs as local anchors is different from that in the studies of fronted objects because there are large categories of PPs that can be shown to be discourse-old on the basis of form. This section therefore starts by selecting these PPs, before analysing the information status PPs that do not contain a discourse-old element.

Let us first consider the overall frequency of clause-initial PPs in *Orosius* and *Lives of Saints*, and compare these numbers to the results for the entire YCOE. Table 4.22 gives the results of Query 4.4, selecting clause-initial PPs in both conjunct and non-conjunct main clauses. The results for *Orosius* and *Lives of Saints* are presented together with the results for the subperiod in which they occur (O2 and O3, respectively), and the final row shows the results for the entire YCOE.

Table 4.22 shows that there is a striking difference between the two selected texts, 18.5% for *Orosius* against 3.6% for *Lives of Saints*. The data for O2 and O3 indicate that this correlates with a difference between the subperiods from which the texts are selected, although the gap for the entire periods is smaller, with 9.3% in O2 and 6.2% in O3. The percentage for *Orosius* is considerably higher than the overall percentage for the entire YCOE, with 18.5% against 6.2%, while the percentage for *Lives of Saints* is lower than this, with 3.6%.

Table 4.22 Percentage of clause-initial PPs in *Orosius*, *Lives of Saints* and the entire YCOE

	No. of clause-initial PPs	Total no. of main clauses with Sbj and FinV	% clause-initial PPs
<i>Orosius</i>	380	2056	18.5
O2	692	7477	9.3
<i>Lives of Saints</i>	153	4216	3.6
O3	1193	21108	5.5
Entire YCOE	3198	51630	6.2

Between *Orosius* and *Lives of Saints*: $\chi^2(1)$ 392.12, $p = 0$

It seems, then, that both *Orosius* and *Lives of Saints* are quite exceptional in the frequency with which they use clause-initial PPs, which is possibly due to the specific genres they belong to, or an idiosyncrasy of these two texts: we have already seen that Ælfric's work is highly stylistic, while *Orosius* contains stretches which are less narrative and more formulaic, which may skew the numbers. While this means these texts are not as representative as would be desirable, these extremes could in fact provide interesting data with respect to use of clause-initial PPs.

Table 4.23 presents a breakdown of the clause-initial PPs according to their anaphoric elements (demonstratives and pronouns), based on Queries 4.4 and 4.5: 'P+D' and 'P+PRO' for NP complements with only a demonstrative or pronoun; 'P+NP(D/PRO)' for NP complements which contain a demonstrative or pronoun and additional material; and 'No D/PRO' for NP complements which do not contain a demonstrative or pronoun.

Table 4.23 Anaphoric elements in clause-initial PPs in *Orosius* and *Lives of Saints*⁶¹

Type of PP	<i>Orosius</i>		<i>Lives of Saints</i>		Both texts	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
P+D	175	46.1	23	15.0	198	37.1
P+PRO	14	3.7	4	2.6	18	3.4
P+NP(D/PRO)	164	43.2	90	58.8	254	47.7
No D/PRO	27	7.1	36	23.5	63	11.8
Total	380	100.0	153	100.0	533	100.0

Between *Orosius* and *Lives of Saints*: $\chi^2(3)$ 59.14, $p = 0$

⁶¹ There were 4 examples of 'No D/PRO' in *Orosius*, and 8 in *Lives of Saints* which did in fact contain a demonstrative or pronoun. They were added to the P+NP(D/PRO) category.

Table 4.23 shows that roughly 40% of the clause-initial PPs in the two texts are in all likelihood old: they have a demonstrative (37.1%) or a pronoun (3.4%) as the only element in the complement position of the P, as in (118).

- (118) =(110)
Mid þam wunode an mæden mærllice drohtnigende,
 ‘With him [lit. that] dwelt a virgin, living virtuously,’
 (coelive [Basil]:468.786)

Similar to the previous table, there is a considerable difference between the two texts, with ‘P+D’ making up almost half of the PPs in *Orosius* (46.1%), while only representing 15.0% of the examples in *Lives of Saints*. This means that there is not only a quantitative difference between the two texts, but also a qualitative one. It is not clear, at this point, what could explain this difference; I will return to this issue below.

There are an additional 47.7% of clause-initial PPs which also contain an anaphoric element, but which contain additional material. These NPs will in all likelihood be OLD or ACCESSIBLE, because they contain an anaphoric element, but exceptions are certainly possible (cf. PDE *I met this guy yesterday*, where *this* is used even though a new referent is introduced). Let us look at this category in more detail, to determine whether classifying these NPs as discourse-old is justified. The PPs come in two categories. The first category of examples contains a specific reference to an antecedent in the preceding clause, as in 0 and (120). The antecedents are underlined.

- (119) *Ælc cristen man sceal cunnan his paternoster and his credan.*
 every Christian man must know his Pater-Noster, and his Creed
Mid þam paternostre he sceal hine gebiddan, and mid þam credan he sceal his
geleafan getrymman.
 with the paternoster he shall to-him pray, and with the Creed he shall his faith
 confirm
 (coelive [Ash_Wed]:262.2850-2)
- (120) *& þonne be norþan þæm Gandes muþan, þær þær Caucasus se beorg endað*
neh þæm garsecge, þær is se port Samera.
 and then to the-north of-the Ganges’s mouth, there where Caucasus the
 mountain ends, near the ocean, there is the port Samera
Be norþan þæm porte is se muþa þære ie þe mon nemneð Ottorogorre.
 to the-north of-the port is the mouth of-the river which they call Ottorogorre
 (coorosiu,Or_1:1.9.27.123-4)

In 0, the NPs *Mid þam paternoster* and *Mid þam credan* have antecedents in the immediately preceding clause: *his paternoster* and *his credan*. The same holds for

example (120), where *Be norþan þæm porte* refers back to *se port Samera* in the preceding sentence. In both these examples, there is a specific antecedent in the preceding clause and part of this NP is repeated in the clause-initial PP.

The second type of PP consisting of a P and an NP containing a demonstrative has a more general reference to preceding events: the reference is to the immediately preceding sentence, but there is no specific antecedent, as illustrated in (121)-(123); cf. the event, topic, and speech categories for objects and PPs.

(121) ‘Then the boy told it to the sacristan, and said, that he had never before been able to speak, praying that they would sing the appointed hymn of praise.’

On þære ylcan tide wæs sum wyn gehæft on swinglum for swyðe lytlan gylte,
about the same time was a-certain bondwoman caught to be-flogged for [a]
very slight fault

(coelive [Swithun]:166.4316-8)

(122) =(112)

“Clearly thou knowest that He is God’s very Son, Who gave to the dead life after death, and to sick men healing by His Word.”

Æfter ðysum wordum se eadiga Petrus gehadode Apollonarem and gehalgode to bisceope,

after these words the blessed Peter ordained Apollinaris and consecrated [him] as bishop

(coelive [Apollinaris]:25.4540-1)

(123) ‘For that fault, his father then ordered him to be put to death: because of that death, the Romans would not, as was their custom, offer the triumph to the consul, though he had gained the victory.’

On ðæm æfterran geare þæs, Minucia hatte an wifmon, þe on heora wisan sceolde nunne beon.

in the following year from-that, Minucia was-called a woman, who in their manner, should a-nun be

‘In the following year, there was a woman called Minucia, who according to their custom should be a nun’

(coorosiu,Or_3:6.60.8.1153-4)

In (121), *On þære ylcan tide* refers to the time at which the events discussed in the preceding section took place, but there is no specific reference. In (122), *Æfter ðysum wordum* refers to the act of speaking, but does not refer back to specific ‘words’. And (123), again, has a clause-initial PP *On ðæm æfterran geare þæs* that only very generally refers back to the preceding sentence. It is, of course, possible that there are PPs with an anchored element which is not necessarily old, so this is a clearly different category from the P+D group. Nevertheless, it seems safe to conclude that the vast majority of these are in fact discourse-old, which means that

the total percentage of discourse-old elements in the two texts may be as high as 87.4%.

The final result from Table 4.23 is that only 11.8% (63 examples) of the clause-initial PPs in *Orosius* and *Lives of Saints* are not clearly discourse-linked on the basis of their form. Again, the two texts behave differently: where *Orosius* only has 7.1% of the examples with no indication of a discourse link, in *Lives of Saints*, they make up roughly a quarter of the examples (23.5%). These PPs belonging to the ‘No D/PRO’ category were manually analysed for information status and anaphoricity, with the exception of 4 examples whose information status could not be analysed.⁶² Table 4.24 shows the results for information status.

Table 4.24 Information status of clause-initial PPs with no anaphoric element in *Orosius* and *Lives of Saints*

Information status	<i>n</i>	%
OLD	16	27.1
ACCESSIBLE	6	10.2
ANCHORED	18	30.5
NEW	19	32.2
Total	59	100.0

$\chi^2(3) = 7.24, p = 0.06$

The table shows that just over a third of the examples in the rest category are ‘given’, i.e. OLD or ACCESSIBLE (37.3%, 22 of the 59 examples), while the other two-thirds are ‘new’, i.e. ANCHORED or NEW (62.7%, 37 out of 59). This means that the majority of examples represent cognitively new referents. If we consider the discourse behaviour of these elements, however, the picture changes slightly, because two-thirds of the examples are in fact linked to the discourse (the OLD, ACCESSIBLE and ANCHORED categories). In other words, a distinction can be made between cognitive accessibility and discourse behaviour, and it may be that discourse behaviour is more important for initial PPs than cognitive accessibility.

The following example illustrate an ACCESSIBLE (124) and a NEW (125) PP.

(124) =(114)

On Easterdagum he wolde etan fisc gif he hæfde.

‘On Easter-days he would eat fish if he had it.’

(coalive [Martin]:1267.6800)

(125) *ON sumere tide he ferde forð þurh ane burh Ambianis gehaten,*

on a-certain occasion he travelled forward through a town, Amiens called

(coalive [Martin]:58.6001)

⁶² One PP was in fact a clause, one was an adverb, and two contained negative NPs and were therefore excluded.

These two PPs seem to be frame setters in the sense of Krifka (2007), who defines frame setters as clause-initial elements which provide the general context for a following statement, but which also provide the restrictions within which to interpret the information (Krifka 2007: 36, who reconsiders an earlier definition given by Chafe 1976). The relevant aspect of Krifka's definition is his notion of alternatives, which frame-setters share with contrastive elements. The distinction between contrast and frame setting that I will use throughout this thesis is based on the presence or absence of alternatives in the text. This is not a necessary condition for Krifka, who works with modern languages, but it is a necessary condition for diachronic work. Whenever there is an alternative present in the text, the example will be analysed as contrastive, while the contrast can remain implied for frame setters.

Another aspect of the information-structural behaviour of clause-initial PPs in OE is the presence of and distance to an antecedent. Table 4.25 presents the results for the examples belonging to the 'No D/PRO' category. The category '1 clause' includes references to a larger set, while '1 clause (as anchor)' includes inferred links, anchors and alternatives.

Table 4.25 Anaphoricity of clause-initial PPs with no anaphoric element in *Orosius and Lives of Saints*

Distance to antecedent	<i>n</i>	%
1 clause	10	16.9
1 clause (as anchor)	17	28.8
2-5 clauses	6	10.2
2-5 clauses (as anchor)	2	3.4
No antecedent	24	40.7
Total	59	100.0

$\chi^2(4) = 26.17, p < .001$

The table shows that almost half of these PPs (40.7%) are not linked to the preceding discourse, and only a small number (16.9%) have an explicit antecedent. Combining these scores with the categories of PP containing demonstratives and pronouns, which I will take to refer to the preceding sentence, means that in total only a very small number of the clause-initial PPs do not have an antecedent in the preceding 5 sentences, again showing that discourse behaviour is an important factor, possibly separate from information status. Interestingly, a considerable number of the 59 examples (27.1%, 16 examples) are contrastive, illustrated in (126) and (127).

(126) 'Now will we speak about Greece, on the south of the river Danube. The sea, Propontis, lies on the east of Constantinople, a city of the Greeks.

& be norðan Constantinopolim Creca byrig scyt se sæearm up of þæm sæ westrihte þe man hætt Euxinus;

and on the.north of.Constantinople of.the.Greeks city shoots the sea-arm up from the sea western that they call Euxine

‘and on the north of Constantinople, a city of the Greeks, the arm of the sea shoots up right west from the Euxine;’

(coorosiu,Or_1:1.18.5.349)

- (127) *& on Tracio þæm londre mon geseah swelce se heofen burne; & on Ariminio þære byrg wæs niht oð midne dæg.*

and on Thrace the country they saw as-if the heaven burnt; and on Ariminum the city was night till mid day

‘and in the country of Thrace, they saw, as if the heaven were burning; and in the city, Ariminum, it was night till mid-day’

(coorosiu,Or_4:7.98.27.2030-1)

In (126), *be norðan Constantinopolim* is presented in contrast to ‘on the east of Constantinople’ in the previous sentence. In (127), the two locations of *Tracio* and *Arminio* are contrasted. The alternatives with which the PPs are contrasted are in both examples explicitly present in the text. Note, however, that information status or anaphoricity does not directly correlate with markedness (contrast or frame-setting); it possible for a discourse-old element to be contrastive.

This study of the clause-initial PPs in *Orosius* and *Lives of Saints* has shown that a large part of the PPs in clause-initial position are discourse-old and contain a link to the preceding discourse, in most cases to the immediately preceding sentence. In other words, clause-initial PPs—like clause-initial objects—are used to express old information and to provide a link to the immediately preceding discourse. In addition to this function of local anchoring, there are other, less frequent but clearly relevant patterns—frame setting and contrast—which indicate the versatility in information-structural categories of clause-initial PPs in OE and which show that the patterns that are dominant in PDE were already present in OE, but as minority patterns only.

4.4.3 Inversion in PP-initial clauses

The second issue which is relevant to the discussion of clause-initial PPs is the question whether information-structural factors determine inversion of the subject and the finite verb. This question is slightly different for clause-initial PPs than for objects, because the variation in syntactic and semantic categories of PPs and the lack of data in the literature on the effect of these different categories on inversion (cf. the differentiation for adverbs, with some always triggering inversion) means that we first need to consider other factors before we can determine the influence of information status. This section aims to establish whether the type of PP and the type of verb have an influence on inversion, considering the entire YCOE, but with a special focus on the periods in which *Orosius* and *Lives of Saints* occur, O2 and O3.

Table 4.26 shows the rates of inversion with nominal subjects after clause-initial PPs throughout the entire YCOE, including both conjunct and non-conjunct clauses. The table includes the PP-initial clauses of the selection in the previous section, but now only with nominal subjects.⁶³ Recall the three subject positions in OE: there is not just a variation between V2 and V3 surface word order, but there is also the late subject position within the VP to consider. While late subjects are unlikely to occur in object-initial clauses, which are transitive, for PP-initial clauses it is important to attempt to distinguish between immediately postverbal and late subjects, or at the very least, take into account that subjects that occur after the finite verb are not all the result of V2 inversion, but occur lower in the clause for other syntactic and information-structural reasons. Nevertheless, the queries used here do not initially distinguish between inverted (Spec,TP) and late subjects (within the VP), but only only between preverbal (Spec,FP) and any postverbal subject. The reason for this is the difficulty in determining the position of the postverbal subject because diagnostic elements (such as non-finite verbs) may not always be available. Table 4.26, then, aims to provide a general indication of the number of postverbal subjects to provide an initial view on inversion in the database.

Table 4.26 Inversion of nominal subjects in PP-initial main clauses in the YCOE (conjunct and non-conjunct clauses combined)

	Inversion	Total	% Inversion
<i>Orosius</i>	153	316	48.4
O2	286	524	54.6
<i>Lives of Saints</i>	95	110	86.4
O3	696	816	85.3
Entire YCOE	1689	2344	72.1

Between *Orosius* and *Lives of Saints*: $\chi^2(1) = 48.3$, $p = 0$; Between O2 and O3: $\chi^2(1) = 153.75$, $p = 0$

The table shows that there is once again a difference between the two texts, similar to what we already saw for inversion in object-initial clauses: while *Orosius* shows inversion in almost half of the PP-initial clauses (48.4%), inversion is clearly a majority pattern in *Lives of Saints*, with 86.4%. Again, this seems to be a reflection of the overall difference between the two periods to which these texts belong, with 54.6% inversion in O2 against 85.3% inversion in O3. The average for the entire YCOE lies in between these percentages for O2 and O3, at 72.1%.

Although it is not possible to select all late subjects, it is possible to select some of them. The total number of late subjects in passive clauses—subjects which follow the participle—is 102 in the entire YCOE, while in *Orosius* there is one and in *Lives of Saints* there are seven. The overall percentage of inversion for the entire YCOE goes

⁶³ A query was added to those in Query 4.4, which stipulated that the finite verb should precede the subject (inversion), or the subject should precede the finite verb (non-inversion). An additional query was added (subject idomonly *PRO*) to select and then exclude the pronominal subjects.

down from 72.1% to 70.8% if those subjects are excluded, which can hardly be said to be a major influence. Another way of selecting possible late subjects, by selecting clauses with other material intervening between the finite verb and the subject, also does not help much: there are only a handful such subjects in *Orosius* and *Lives of Saints* (15 and 17, respectively). In addition, while an intervening XP might indicate a late subject, it does not mean that all these sentences have a late subject (unlike post-participle subjects).

Before we look at inversion more closely, let us consider the difference between non-conjunct and conjunct clauses, one of the factors said to influence inversion in OE. Table 4.27 presents the rates of inversion in conjunct and non-conjunct clauses in O2, O3 and the entire YCOE.

Table 4.27 Inversion of nominal subjects in PP-initial main clauses in the YCOE (conjunct and non-conjunct clauses compared)

	Non-conjunct			Conjunct		
	Inversion	Total	%	Inversion	Total	%
O2	239	449	53.0	48	75	64.0
O3	590	669	88.2	106	147	72.1
Entire YCOE	1319	1784	73.9	370	561	66.0

Entire YCOE, between non-conjunct and conjunct: $\chi^2(1) = 13.49, p < .001$

The table shows that the inversion rates for non-conjunct clauses are roughly similar to those found in the previous table (just over 50% for O2, over 80% for O3, and over 70% for the entire YCOE), but the percentages of inversion in conjunct clauses are decidedly different: for O2, O3 and the entire YCOE, the percentages are in the region of 70%. They still reflect the percentages for non-conjunct clauses—lower for O2 and higher for O3—but on a smaller scale. Most importantly, the rates of inversion seem to be more constant throughout all periods, although the exact reason for this would need a separate investigation. Because conjunct clauses behave differently, for reasons that are not clear, the rest of this section focuses on non-conjunct clauses only.

In later stages of the language, several factors have been shown to influence inversion, often specifically in PP-initial clauses—see e.g. Van Kemenade (1997) and Warner (2007) for Middle English, and Bolinger (1977) and Birner & Ward (1998) for PDE. The question is whether these factors already influence inversion in OE and should therefore be excluded from a study of the information status of the subject. The first of these factors is the type of verb, with unaccusatives showing more inversion than transitive verbs. As an approximation of unaccusative verbs, Table 4.28 shows the rates of inversion in clauses where *be* is a finite verb, and there may or may not be an additional non-finite verb in the clause. In most cases the non-

finite verb will be the passive participle, which means these clauses can justifiably be grouped with unaccusative verbs.⁶⁴

Table 4.28 Inversion with nominal subjects in clauses with *be* as the finite verb

	Inversion	Total	% Inversion
O2	92	119	77.3
O3	233	244	95.5
Entire YCOE	678	749	90.5

Between O2 and O3: $\chi^2(1) = 28.21$, $p < .001$

The table very clearly shows the effect of the unaccusative verb *be*, most clearly in the much higher percentage of inversion in O2 compared to the overall rate of inversion (77.3% against 53.0%, see Table 4.27); even in O3 and the entire YCOE, where the percentages were already quite high, an effect is visible, with the percentages reaching over 90 (95.5% against 88.2% for O3, 90.5% against 73.9% for all periods).

Another possibility for investigating the influence of unaccusative verbs is to select transitive main clauses, i.e. clauses which are clearly not unaccusative, by selecting clauses which have both a subject and direct object, and to compare the rate of inversion in those sentences against the overall rate of inversion.⁶⁵ These sentences only represent one category of transitive clauses because adjectival and prepositional complements are not included (as this cannot be done on the basis of form either), but they are prototypical transitive verbs and as such are a reliable category to contrast with the findings for *be*. Table 4.29 presents the results.

Table 4.29 Inversion after PPs in clauses with a subject and an object in main clauses

	Inversion	Total	% Inversion
O2	13	78	16.7
O3	64	87	73.6
Entire YCOE	95	239	39.7

Between O2 and O3: $\chi^2(1) = 53.49$, $p = 0$

The first thing to note in this table is the small number of examples: only 239 in the entire YCOE (compare the number of examples in the previous table: 749 clauses with *be*). The reason for this is partly that the table represents only a limited selection of transitive clauses and it excludes not only clauses with prepositional complements but also clauses with ellipted subjects. Generally, however, it seems to

⁶⁴ The labels that were used for finite *be* are (*BED*|*BEP*), and a query selecting these labels was used in combination with the queries selecting non-conjunct clauses and inversion/non-inversion.

⁶⁵ The queries that were used were similar to those for transitive clauses in section 4.3, so a subject (NP-NOM|NP-NOM-RSP|NP-NOM-1) and object (NP-ACC|NP-ACC-RSP|NP-ACC-RFL|NP-ACC-1|NP-ACC-RSP-1|NP-ACC-RFL-1) as main clause constituents. These queries were combined with the other queries used here to select PP-initial clauses, and inversion vs non-inversion.

be the case that transitive clauses do not often have a PP in clause-initial position. Nevertheless, the table confirms the expectations about transitive sentences: the rates of inversion in this table are lower than the average rates of inversion in the selected texts. For O2, the percentage is particularly low: only 16.7% inversion, against 53.0% inversion on average and 77.3% in clauses with *be*. These two simple ways of comparing unaccusative verbs with transitive verbs, then, confirm the general tendency that unaccusative verbs trigger inversion to a higher degree than other types of verbs.

The second factor shown to influence inversion in later periods is the type of clause-initial PP, with locative PPs apparently acting as a trigger for inversion. Unfortunately, information about the type of PP—in terms of location, time, manner or means—is not available in the YCOE, but what is possible is to select the PP on the basis on its preposition. This is not completely reliable since one preposition can express different functions, with *on* encoding both location (*on the beach*) and time (*on that day*). Nevertheless, the first step, and the one that we can make here, is just to see whether the choice of preposition makes a difference at all. The following table presents the rates of inversion after a selection of prepositions. The numbers in between brackets after the percentage give the number of inverted subjects and the total number of PPs with this preposition.

Table 4.30 Inversion with nominal subjects in PP-initial main clauses, differentiated on the basis of type of P⁶⁶

	<i>On</i>	<i>After</i>	<i>Through</i>	<i>In</i>	<i>To</i>	<i>With</i>
O2	58.2% (71/122)	28.9% (37/128)	60.0% (3/5)	55.6% (15/27)	50.0% (3/6)	66.7% (6/9)
O3	92.0% (243/264)	79.2% (80/101)	93.1% (27/29)	0.0% (0/1)	88.0% (22/25)	100.0% (2/2)
Entire	83.5% (128/775)	52.9% (135/256)	73.1% (38/52)	51.0% (25/49)	65.1% (28/43)	48.1% (13/27)

O2: $\chi^2(5) = 25.35$, $p < .001$; O3 (without *in* and *with*): $\chi^2(3) = 12.68$, $p < .01$; Entire YCOE: $\chi^2(5) = 213.55$, $p = 0$

The first thing to note from this table is that the number of examples differs considerably between prepositions and between periods, ranging from 775 examples of *on* in the entire YCOE to 27 examples of *with*. The table provides evidence that the choice of preposition does indeed make a difference for inversion, but at the same it is difficult to determine the exact influence across different texts. If we consider O2, for instance, inversion is less common with *after* (28.9%) compared to the other prepositions, and slightly more frequent than the average

⁶⁶ All relevant spelling variants of each preposition were included, based on a 'lexicon' search by *CorpusSearch*.

from Table 4.25 in clauses with *on*, *through* and *with* (57.9%, 60.0% and 66.7% respectively, against 53.0%). Similarly, for O3, inversion is less frequent with *after* (79.2% against 87.4% for all PPs, disregarding *in* which has 0.0% but based on only a single example), but roughly equally frequent as the average with *to* (88.0%) and slightly more frequent with *on* (92.0%) and *through* (93.1%). In short, the range seems to be wider for O2 and the entire YCOE (from 51.0% to 83.5%), while in O3 inversion rates are relatively stable across all prepositions.

A final way to approach the question of variation after different types of PPs is to categorize them according to their information status—relevant in light of the proposed information-rearranging function of locative inversion (see Birner & Ward 1998). The following table presents the rates of inversion, using the distinctions for information status made previously on the basis of the form of the PPs.

Table 4.31 Inversion with nominal subjects after different PP categories in main clauses

	P+D	P+PRO	P+NP(D/PRO)	No D/PRO
O2	43.8% (92, 210)	100.0% (5,5)	58.1% (111, 191)	69.8% (30, 43)
O3	92.1% (140, 152)	62.5% (5,8)	90.6% (298, 329)	81.7% (147,180)
Entire YCOE	61.4% (290,472)	80.0% (12,15)	80.8% (780, 965)	71.4% (237,332)

O2 (without P+PRO): $\chi^2(2) = 13.92$, $p < .001$; O3 (without P+PRO): $\chi^2(2) = 16.32$, $p < .001$; Entire YCOE: $\chi^2(3) = 63.44$, $p = 0$

If we compare the rates of inversion in this table to the rates of inversion in Table 4.26, it turns out that in O2 and the entire YCOE, inversion is less frequent than on average after clause-initial PPs that only consist of a preposition and a demonstrative, such as *after that* (43.8% for O2 against 53.0% average, 61.4% for entire YCOE against 73.9% for all PPs). The O3 period does not show this: 92.1% in this table, 87.4% for all PPs). In contrast, inversion is more frequent than average after PPs consisting of a preposition followed by a noun containing a demonstrative or pronoun (*in those days*): 58.1% for O2 (against 53.0%), 90.6% for O3 (against 87.4%) and 80.8% for the entire YCOE (against 73.9%). There is no immediate explanation of why the form of the PP might determine inversion. It does seem to be the case that P+D is a specific pattern and perhaps these clauses favour inversion for a prosodic reason: the clause-initial element is short (cf. Speyer 2010, who also links inversion to prosody). A related consideration is the fact that most of these grammaticalize in later stages into adverbs—as an example, *after þam* is consistently translated as ‘afterwards’ in *Orosius* (Bosworth 1859)—and it could be that speakers already perceive them as one word, an adverb, rather than a full preposition and pronoun at this period, especially in those cases where the pronoun does not refer

back to a specific antecedent but only provides a general reference. In regular PPs it is the NP which provides the link to the previous discourse. Once the P and NP have merged into one adverb, the specificity of referring is lost; adverbs can only make general references to the previous discourse.

In conclusion, this section has identified several factors that influence inversion in PP-initial clauses in OE—based on factors shown to influence inversion in PDE—in addition to the difference between the O2 and O3 subperiods that we previously identified. Surprisingly, conjunct clauses showed a rather constant rate of inversion throughout all texts, with hardly any variation; in contrast to what has been said previously in the literature this did not always mean a lower rate of inversion. As an approximation of unaccusative verbs, I showed that clauses with *be* clearly preferred inversion and straightforwardly transitive clauses (with a subject and a direct object) disfavoured inversion, confirming what has earlier been proposed for later stages of the language. The final two studies aimed to establish the influence of the type of PP and showed that the type of PP does make a difference, but it is difficult to explain why such a difference might exist, for instance in the case of P+D PPs disavouring inversion. What this section has especially made clear is that within the category of PPs, there is considerable variation between types and properties of the PPs in triggering inversion—more so than for object-initial clauses—and it remains to be seen how much room is left for information-structural factors to play a role.

4.4.4 Information-structural factors in *Orosius*

This section aims to establish the relevance of information-structural factors in the position of subjects in pre- and postverbal position in PP-initial clauses. Like for the subjects in object-initial clauses, the question is whether there is a relation between the position of the subject and its information status: in short, are preverbal subjects old and postverbal subjects new? In contrast to the previous section, I do not attempt to distinguish between V2 inversion and late subjects, but only consider the difference between preverbal and postverbal subjects. The reason for this is, in addition to the problems of determining the position of the subject, that the predictions for the information status of these subjects is similar: they are both expected to be new (see Biberauer & Van Kemenade 2011, Hinterhölzl & Petrova 2010).

First of all, the table below presents the rates of inversion in non-conjunct clauses in *Lives of Saints* and *Orosius*.

Table 4.32 Inversion with nominal subjects in non-conjunct, PP-initial clauses

	Inversion	Total	% Inversion
<i>Orosius</i>	119	267	44.6
<i>Lives of Saints</i>	78	89	87.6

$\chi^2(1) = 50.10, p = 0$

Not surprisingly, the rate of inversion in *Lives of Saints* is high, 87.6%, while in *Orosius*, inversion and non-inversion are more or less equally frequent, with 44.6% inversion. The total number of examples is also lower in *Lives of Saints* than in *Orosius* (89 against 267), which, together with the high percentage of inversion, leads to a very low number of uninverted subjects (11). This number is too low to allow us to find out whether other factors, such as information status, play a role. For this reason, I focus in this section on the data from *Orosius*, where inversion is more variable, so that the influence of information-structural factors can be expected to be stronger. However, based on the discussion in the previous section, a number of clauses were excluded from the study because they contain a PP or verb which clearly influences inversion. First, clauses with a finite form of *to be* were excluded because they clearly prefer inversion. Second, clauses with a PP consisting only of a preposition and a determiner were also excluded, because they favour uninverted order, as shown in the following table.

Table 4.33 Inversion with nominal subjects after different PP categories in main clauses in *Orosius*

	Inversion	Total	% Inversion
P+D	30	128	23.4
P+PRO	4	4	100.0
P+NP(D/PRO)	77	127	60.6
No D/PRO	8	8	100.0
Totals	119	267	44.6

$\chi^2(3) = 51.32, p = 0$; without 'P+PRO' and 'none': $\chi^2(1) = 36.21, p = 0$

When the clauses with *be* and the clauses with 'P+D' in initial position are excluded, 77 subjects remain, with 49.4% inversion. The information status of these nominal subjects was analysed and the results are shown in Table 4.34.

Table 4.34 Information status of nominal subjects in PP-initial main clauses in inversion and non-inversion contexts in *Lives of Saints* and *Orosius*, excluding PPs consisting of only P+D

	Inversion		Non-inversion		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
OLD	15	37.5	25	62.5	40	100.0
ACCESSIBLE	0	0.0	1	100.0	1	100.0
Subtotal 'Given'	15	36.6	26	63.4	41	100.0
ANCHORED	15	55.6	12	44.4	27	100.0
NEW	8	88.9	1	11.1	9	100.0
Subtotal 'New'	23	63.9	13	36.1	36	100.0
Total	38	49.4	39	50.6	77	100.0

$\chi^2(3) = 9.27, p < .5$; between 'Given' and 'New': $\chi^2(1) = 5.72, p < .5$

NEW subjects have a clear preference for inversion (88.9% against 11.1%), while OLD subjects have a preference for uninverted order, although the preference is less pronounced (62.5% non-inversion against 37.5% inversion). The ANCHORED subjects do not show a preference—48.3% inversion—while the database only contains one example of an ACCESSIBLE subject. If we consider only the binary distinction, the table shows that there is a roughly equal preference for inversion with ‘new’ subjects (63.9%) and for non-inversion with ‘given’ subjects (63.4%). These figures are in line with the general prediction that OLD subjects generally occur before the finite verb, while NEW subjects follow it, although there are quite a number of exceptions, especially for the OLD subjects. Interestingly, the preference for ‘given’ subjects to be uninverted (63.4%) is higher than it was for objects (49.0%), while the preference for ‘new’ subjects to be inverted is quite a bit lower (63.9% against 81.3%). In all cases, however, there are quite a number of exceptions and the percentages, while representing majorities, are not extremely high. Finally, if we consider the distribution of information status categories, it turns out that the overall percentage of OLD subjects is relatively low, with 53.2% (41 out of 77), against 88.9% for subjects in passives and 62.7% for subjects in object-initial clauses. In contrast, the category of ANCHORED subjects—a category which is particularly ambiguous in terms of information status and discourse behaviour—is much larger than it was in earlier studies, at 35.1% (27 out of 77), against 3.2% in passives and 13.4% in object-initial clauses.

The most exceptional categories among these subjects are OLD or ACCESSIBLE subjects that are inverted (13), and ANCHORED or NEW subjects that are not inverted (16). Let us first consider the non-inverted ANCHORED and NEW subjects. The great majority (13 of 16) of these subjects are proper names, and they often contain an apposition or another form of postmodification, as illustrated in (128)–(130). The subjects are underlined.

(128) *On þæm dagum Serius Galua, Scipian qefera, feaht wið Lusitaniam, Ispania folce,*

‘In those days, Servius Galba, a colleague of Scipio, fought against the Lusitanians, a people of Spain,’
(coorosiu,Or_4:12.111.6.2322)

(129) *Æfter þæm Pompeius se consul for on Numentinas, Ispania þeode,*

‘Afterwards Pompeius, the consul, marched upon the Numantines, a people of Spain’
(coorosiu,Or_5:2.115.22.2415)

(130) *On þære tide Mitridatis, Partha cyning, geeode Babyloniam,*

‘At that time, Mithridates, king of Pontus, overcame Babylon’
(coorosiu,Or_5:2.115.31.2422)

In all three examples the subject is a proper name—*Serius Galua*, *Pompeius* and *Mitridatis*—which is followed by an explanation of the function of this person: *Scipian gefera*, *se consul* and *Partha cyning*. Most of the names did not occur earlier in the text. What makes the information status of these subjects slightly problematic is that it is not clear whether the names should be considered as part of world knowledge (for instance when the subject is *Petrus se apostol*) or whether their anchoring, such as ‘the king of Pontus’, in fact makes them accessible enough to act as an OLD/ACCESSIBLE subject, i.e. occur in front of the postverbal verb. From that perspective, the name is the additional information and the postmodification is the proper subject. Unfortunately, the behaviour of these ANCHORED subjects in both inversion and non-inversion sentences does not make this any clearer: they occur just as frequently in inverted sentences (14) as they do in uninverted sentences (15).

The OLD subjects present a problem of their own, with a considerable number of old subjects occurring in inverted position, i.e. after the finite verb. What is interesting in this respect is the difference between anaphoricity and information status: while the analysis of the information status OLD takes into account the entire text, anaphoricity only takes into account the five preceding clauses. Consider the following example.

- (131) *On þæm dagum gecuron Brettanie Maximianus him to casere ofer his willan.*
in those days chose Britons Maximus them to emperor against his will
‘In those days the Britons chose Maximus for their emperor, against his will.’
(coorosiu,Or_6:35.153.28.3260)

The subject *Brettanie* occurs postverbally, and is analysed as OLD because it occurs earlier in the text. It does not, however, have a score for anaphoricity, because the previous mention is in *Orosius*’ book 1, while example (131) occurs in book 6. The question is whether something that is mentioned so far back can still be considered discourse-old, or should for instance be understood as world knowledge rather than discourse-old knowledge, in which case it needs to be reactivated here. Recall from the discussion of subject placement in object fronting sentences (section 4.3.4) the observation that placing a subject later in the sentence, i.e. after the finite verb, may be a type of topic introduction, a function that does not stipulate an element of a specific information status: a referent can be newly introduced as topic even if it represents accessible knowledge. The anaphoricity scores, presented in Table 4.35 for all subjects, do not confirm this, however.

Table 4.35 Anaphoricity scores for inverted and uninverted nominal subjects.

Distance to antecedent	Inversion		Non-inversion	
	<i>n</i>	%	<i>n</i>	%
1 clause	3	7.9	4	10.3
1 clause (anchor)	2	5.3	1	2.6
2-5 clauses	4	10.5	4	10.3
2-5 clauses (anchor)	3	7.9	3	7.7
No antecedent	26	68.4	27	69.2
Total	38	100.0	39	100.0

The table shows that the percentage of subjects which have ‘no antecedent’ in the discourse is roughly similar for inverted subjects as for uninverted subjects: 68.4% and 69.2%. In other words, anaphoricity does not determine the position of the subject. This also means that for both inverted and uninverted subjects, the percentage which has an antecedent in the preceding five clauses is low. What this suggests is that while the initial PPs are overwhelmingly ‘given’ and discourse-old, nominal subjects—both before and after the finite verb—are considerably less often ‘given’. This stresses once again the importance of that first position as a local anchoring position, and indicates a relative ordering of elements according to a given-before-new order in the clause.

The third and final consideration is once again the type of verb, which we have already seen is especially relevant to the PP-initial clauses. Among the inversion sentences with an OLD subject, there are indeed a number of typical unaccusative verbs (4 out of 13), illustrated in (132) and (133).

(132) *On þæm morum eardiað Finnas.*
 ‘In these mountains [wastes] dwell Finns’
 (coorosiu,Or_1:1.15.24.269)

(133) *On þære tide gefor Cassander*
 ‘At that time died Cassander’
 (coorosiu,Or_3:11.81.28.1627)

In these examples, then, the reason for the subject occurring after the finite verb (and plausibly remain within the VP) may be due to the fact that the verb is unaccusative, rather than that it is influenced by the information status of the subject. However, the remaining inverted sentences with old subjects contain transitive verbs, as in 0 and (135).

- (134) *On ða wrace fundon Romane ærest þæt hie scipa worhton.*
in the revenge found Romans first that they ships built
'In revenge the Romans first determined to build ships'
(coorosiu,Or_4:6.92.16.1871)
- (135) *On ðæm gefeohte gefliemde Antigonos Umenis,*
in the battle routed Antogonus Umenis
'In the battle, Antigonos routed Umenis'
(coorosiu,Or_3:11.79.15.1567)

As we have seen before, unaccusativity (or the type of verb in general) influences inversion, and sometimes it seems to overrule information-structural considerations, while in other cases the information status seems to be more important. In examples (132)-(135), it may be that additional factors are needed to explain why inversion occurs (e.g. Speyer's (2010) prosodic account).

In conclusion, this analysis of the information status of subjects in a selection of PP-initial clauses in *Orosius* has shown that, similar to object-initial sentences, there is a tendency for OLD and ACCESSIBLE subjects to occur before the finite verb and for ANCHORED and NEW subjects to occur after the finite verb. The tendency is less pronounced, however, than it was for the subjects in object-initial clauses, and the exceptions to the tendencies are not easy to explain, with for instance anaphoricity and anchoring not providing an answer for why ANCHORED or NEW subjects can occur before the finite verb. One factor that does clearly play a role is unaccusativity, with unaccusative verbs preferring inversion, but this factor cannot by itself explain all sentences showing inversion with an OLD subject. While the percentage of inversion in this selection was near to 50%, and so was suitable for an examination of the influence of information-structural factors, it has to be noted that these sentences represent only one selection of all PP-initial clauses in *Orosius*. With respect to information status, the given-before-new tendency reported in the literature and found in section 4.3.4 was confirmed, but remains no more than a tendency.

4.4.5 Summary and conclusions

This section aimed to establish the syntactic and information-structural characteristics of PP-initial clauses in OE. Specifically, it focused on the function of clause-initial PPs, the rates of inversion after PPs and the information-structural factors in subject placement in PP-initial clauses. I showed that clause-initial PPs, like clause-initial objects, perform a function of local anchoring, but that they are also used for contrast and frame setting. PP-initial clauses already show some tendencies towards inversion patterns specified for inversion in later periods, with unaccusative verbs triggering inversion. Finally, subject placement showed a general tendency for given-before-new, but again, with a considerable number of exceptions.

PPs in clause-initial position turned out to be an even more versatile position than was already clear from the fact that they express different functions (adverbials

and arguments). In terms of information-structural behaviour, three categories were distinguished: local anchoring, contrast and frame setting. While the majority of PPs follow a given-before-new order, with 87.4% of the clause-initial PPs being old information, it is also clear that PP preposing is not an alternative to the passive in the same way as fronted objects are. However, the fact that this varied category does show a sensitivity to given-before-new confirms the strength of this information-structural pattern.

With regard to inversion, it became clear that there are a number of factors playing a role in determining the pre- or postverbal position of the verb, more so than for object-initial clauses. One of these factors is the type of verb, with unaccusative verbs triggering inversion to a larger extent, just like they still allow inversion in PDE. While there was no coding of unaccusativity available, different approaches to the type of verb made clear that unaccusative verbs indeed occur more often with inversion in OE. The type of PP also seemed to play a role to some extent, with for instance a clear preference for PPs consisting solely of a preposition and a demonstrative in *Orosius* to occur with an inverted subject. Throughout this chapter, then, clause-initial PPs have been shown to behave differently from clause-initial objects and these tendencies already indicate the foundations for the remnants of inversion with PPs in PDE.

Section 4.5 Conclusion

This chapter investigated the information-structural properties of the passive, object fronting and clause-initial PPs in OE. Specifically, it investigated whether these constructions show an overlap in the function of creating a given-before-new order of information; what the exact information-structural function of the clause-initial position is; and finally, to what extent word order choices are driven by information-structural factors, especially for the subjects in object-initial and PP-initial clauses as showing or not showing inversion. The overall conclusion is that these constructions share a function of creating or maintaining given-before-new order, stressing the relevance of this information-structural tendency across different constructions. It also became clear that the clause-initial position is used by objects and PPs for so-called local anchoring, providing unmarked links to the preceding discourse, but at the same time, in the case of object fronting and PP preposing, for contrastive purposes as well. Finally, while all these sentence types show the relevance of information-structural factors, I argued that they are not part of the syntax.

As for the overlap in function between the constructions, the studies of the long passive and fronted objects showed that the majority of sentences conform to a given-before-new order of information of the first element (the subject in the case of passives, the object in object fronting sentences) and the second element (the agent phrase of the long passive and the subject in object fronting sentence). The majority of PP-initial clauses also follows a given-before-new order of information. There are crucial differences between the three constructions, however, with given-

before-new being the main motivation in long passives, while in object fronting and especially PP-initial clauses it is only one of the functions: both fronted objects and clause-initial PPs occur with contrastive referents, which are in addition higher on the givenness scale than the elements that follow.

All three constructions showed the relevance of discourse linking in clause-initial position, with the majority of fronted objects and clause-initial PPs being old, i.e. having an explicit antecedent in the preceding discourse. In many cases, the reference was to the immediately preceding sentence, indicating that discourse linking is an especially 'local' phenomenon. With respect to the passive, it is clear that subjects are often old and link to the preceding discourse, but this is a property they share with active subjects. A comparison of passive and active subjects showed that passives were not more often clause-initial, i.e. did not more often than active subjects function as discourse linkers first and foremost.

The final aim of this chapter was to establish to what extent the variation that is characteristic of the V2 phenomenon in OE—between V2 and V3 orders in clauses starting with a non-operator—is determined by information-structural factors. The analysis of the information status of the subjects in object-initial clauses and PP-initial confirmed the tendencies reported in the literature that preverbal subjects are usually given and postverbal subjects usually new. There were quite a number of exceptions, however, and I argued that there was not enough support for a formalization of this tendency in terms of the syntactic structure of the clauses.

This chapter has established the role of information-structural factors within the syntactic properties of the OE main clause. Most importantly, it has shown the function of the clause-initial position as one of local anchoring, and the availability of a range of word order options to perform a similar function of creating given-before-new order of information in the sentence. The question that follows from these observations on the interaction between information-structural and syntactic functions in OE is what happens when the V2 constraint is lost. Before we turn to the consequences of the loss of V2, the next chapter first aims to establish what exactly changes in the use of the clause-initial position and the position and use of the subject in ME and EModE, building on what we have seen in this chapter for OE.

5. Aspects of the Loss of Verb Second

Section 5.1 Introduction

This chapter investigates aspects of the loss of verb second (V2) in the Early Modern English (EModE, 1500-1710) and Modern British English (MBE, 1700-1914) periods. It adds to the existing body of literature on the loss of V2, building on the conclusions from the previous chapter where we saw that the clause-initial position had a special function of local anchoring within the Old English (OE) V2 system, and that the variation in subject placement was governed by a general information-structural motivation. The aim of this chapter is to trace what happens to these aspects when V2 is lost, with a special focus on PP-initial clauses. In this chapter, I show that towards the end of the EModE period, the presubject position lost its main OE function of local anchoring in favour of the contrastive and frame-setting uses. At the same time, the position of the subject became more fixed and over the course of the MBE period, the subject became the most frequent, and crucially the only unmarked, way to start a sentence.

We saw in Chapter 2 that the V2 system as it was in place in OE was lost for the largest part during the fifteenth century, while there were certain contexts that allowed inversion of subject and verb for a longer time. The first of these contexts is the type of verb, with unaccusatives showing inversion quite frequently long after the general V2 organization is lost (see Van Kemenade 2012). The second context

that is relevant is the syntactic category of the clause-initial element: Van Kemenade & Westergaard (2012) show that towards the end of the fifteenth century inversion after objects is still at 82.5% (although based on a small number of examples), while inversion after initial adverbs is already down to just over 20% (with unergative verbs and nominal subjects). Similarly, we have already seen that one of the contexts in PDE in which inversion is still found, locative inversion, is to a large extent (but not entirely) defined by both the verb and the clause-initial element: unaccusative verbs and locative PPs. The loss of inversion with this last category of clause-initial elements has been documented in less detail than the loss of inversion after objects and adverbs in the literature, and its continuing pattern of inversion provides a valuable opportunity to add to an understanding of the factors involved in the loss of V2.

What this chapter specifically aims to do is trace the two information-structural functions that were identified in the previous chapter—local anchoring and creating given-before-new order—within the larger context of the loss of V2. We saw in the previous chapter how the functions of the first constituent in OE were decidedly different from its functions in PDE: while most of the clause-initial objects and PPs in OE are non-contrastive and given, the literature on PDE shows that the clause-initial position is generally used to express contrast (see e.g. Krifka 2007; Speyer 2010; Los & Dreschler 2012). What is not yet clear, however, is when exactly, in the intervening centuries, these changes take place, i.e. when the clause-initial position loses its main function of local anchoring and when the contrastive and frame-setting uses become dominant.

The aim of this chapter is to investigate the decline in inversion and the information-structural functions of the clause-initial position, focusing on the most versatile element in initial position, PPs. One important aspect in studying these consequences of the loss of V2 is the timing of the various developments, just as the timing of the introduction of the new passives relative to the loss of V2 was important in Chapter 3. This chapter specifically aims to establish a timing for two different aspects of the loss of V2: on the one hand the position of the subject with respect to the finite verb, and on the other hand the use of the first position by non-subject elements and the changes in its information-structural motivation. It is these changes that lead to the increased importance of the subject as an unmarked theme, in turn increasing the need for passives as subject creators. The studies in this chapter focus on the sixteenth century, the period immediately following the end date usually given for the loss of V2, but include data from other periods where necessary and available.

In this chapter, I show that the largest decline in inversion in PP-initial clauses takes place towards the end of the EModE period and that the main development towards a predominantly subject-initial language takes place somewhat later, in the MBE period. The origins of the locative inversion pattern that remains in use in PDE are already visible from the beginning of the EmodE period onwards, with the most important characteristics of unaccusative verbs and PPs (either locative or local

anchors): inversion in these instances ‘lingers longer’ than inversion in other contexts. The studies in this chapter provide evidence that the local anchoring function of the clause-initial position becomes less important and the contrastive and frame-setting functions become more important, a development which is already visible in EModE and which continues in the MBE period.

The chapter is structured as follows. Section 5.2 investigates how the position of the subject changes and what role information-structural factors play in determining the position of the subject. Section 5.3 provides evidence for the loss of local anchoring as the most important function of the presubject position, focusing again on PP-initial clauses but with an additional discussion of the loss of local anchoring in object-initial clauses. Section 5.4 concludes this chapter.

Section 5.2 The loss of inversion in PP-initial clauses

This section investigates the loss of inversion in PP-initial clauses from the sixteenth until the early twentieth century. It focuses on two aspects: on the one hand the development towards a default position for the subject and on the other hand, the beginnings of the specific type of inversion that still continues to be used in PDE: locative inversion. The aim is to answer the question whether the relation between information-structural factors and syntactic factors has changed since the OE period. I show that this interaction indeed changes as syntactic options become more limited: with the fixation of a default subject position, information-structural factors begin to play a different role, becoming less important for the default subject position but making full use of the remaining syntactic exception, the late subject position.

The loss of V2 is of course not limited to just the development of the subject in PP-initial clauses, but there are several reasons why the category of PP-initial clauses is especially relevant. Not only do PPs show the range of information-structural options that any element in first position can show, they are also a versatile category syntactically, expressing both arguments and adjuncts. In addition, most of the studies dealing with the loss of inversion, such as Van Kemenade (2012) and Haeberli (2002), have focused on adverbs or objects in clause-initial position, and less so on PP-initial clauses, although there are some exceptions, such as Speyer (2010) and Los & Dreschler (2012). Another aspect of PP-initial clauses is that, unlike objects, PPs remain quite frequent in clause-initial position in MBE and PDE, which means they can provide more insight into a change in function.

Inversion in PP-initial clauses, especially when these PPs are locative, is still a productive syntactic pattern in PDE, although not a very frequent one. Loosely referred to as locative inversion, this pattern has been said to be dominated by intransitive (more specifically, unaccusative) verbs and locative PPs in clause-initial position (see Bresnan 1994, Levin & Rappaport Hovav 1995, Birner & Ward 1998). What is also relevant for the present investigation is that this pattern has been described as having an information-structural function, either as information-

rearranging, like the passive, or as a presentational construction (Birner & Ward 1998 and Bolinger 1977, respectively). Already in OE, unaccusatives and passives allowed for the late subject pattern but it is unclear from which time onwards the locative inversion pattern emerges and how it interacts with the loss of V2.

This section is structured as follows. Section 5.2.1 discusses the characteristics of the locative inversion pattern in PDE, aiming to define the pattern in such a way that it will allow for a search for its origins in the historical corpora. Section 5.2.2 then turns to the EModE period, investigating the rate of the loss of inversion in PP-initial clauses as well as the influence of the information status of the subject on its position in the clause. Section 5.2.3, finally, turns to MBE and aims to determine to what extent it still shows old patterns of inversion or whether the locative inversion pattern is already the only available inversion pattern. Section 5.2.4 concludes.

5.2.1 Locative inversion in PDE: effect of PP and verb

Before we turn to the corpora to trace the loss of inversion and the emergence of the locative inversion pattern, let us look in more detail at the literature on inversion in PDE, with a focus on the pattern for locative inversion. The aim of this discussion is to define the phenomenon of locative inversion, so that it can inform the evaluation of instances of surface inversion in the corpora. In this section I show that (i) there are two syntactic aspects that are generally taken to define locative inversion, although they do not hold for all examples: the type of verb and the type of clause-initial element; and (ii) two accounts of the function of locative inversion in PDE in terms information-structural factors have been proposed, but with a slightly different definition.

Locative inversion should be distinguished from other types of inversion in PDE, such as quotation inversion in (1) and subject-auxiliary inversion in (2).

- (1) "I think the hyper-cars are more hype than anything," said Larry Carlat, editor of *Toy And Hobby World*. [*Chicago Tribune*, 11/21/89]

(Birner & Ward 1998: 158, their 202)

- (2) Rarely did I hear such overtones of gratitude as went into the utterance of this compound noun. [=Green 1980, ex. 32e]

(*ibid.* 157, their 199)

Quotation inversion, as in (1), is restricted to sentences with verbs of reporting, and has been shown to be irrelevant to the type of subject-verb inversion which I will be discussing in this chapter (Warner 2007, Birner & Ward 1998: 157-159). Subject auxiliary inversion as in (2) also needs to be excluded as this pattern has a syntactic rather than information-structural trigger, with the verb moving to C in questions and after initial constituents containing negation (including *rarely*) and focus markers such as *only*, *many a* and *such a* (cf. Huddleston & Pullum 2002: 95).

A less well-defined class of inversion, because of its syntactic variability, is what has become known as locative inversion (Coopmans 1989, Bresnan 1994, Culicover 2001). Consider the following sentences, in which the (a) examples illustrate the normal order of subject and finite verb, and the (b) examples illustrate the inverted order.

- (3) a. A lamp was in the corner.
b. In the corner was a lamp.
- (4) a. The tax collector came back to the village.
b. Back to the village came the tax collector.

(Bresnan 1994: 75, her 1,3)

The alternations in the a and b sentences in (3) and (4) show how the subject, with its usual position before the finite verb (*a lamp, the tax collector*), can also occur after the finite verb in those instances where the sentence begins with a prepositional phrase indicating a location, like *in the corner* and *back to the village*. Two central properties have been discussed in the literature in attempts to define the context for locative inversion: the type of verb and the type of element in first position.

First, locative inversion is said to occur only with intransitive verbs (Rochemont 1978), or more specifically, unaccusative intransitives and passivized transitive verbs. The following examples illustrate the possibility for unaccusative (5) and passivized transitive verbs (6) to occur with inversion, and the impossibility of inversion with transitive verbs (7).

- (5) Into the room came an elephant.
- (6) My mother was seated among the guests of honor.

(Bresnan 1994: 78, her 18a)

- (7) *Into the room rolled John the ball.

(Coopmans 1989: 730, his 5a; based on Rochemont 1978)

The presence of passives among the verbs that allow inversion is interesting because of the impossibility for transitive verbs to occur with inversion and the requirement for passivization that a verb be transitive, but not surprising because of what happens in the process of passivization. As Bolinger (1977: 103) phrases it, “[t]he passive makes it possible to use transitive verbs [with inversion] by avoiding the problem of too many arguments. The agent is simply deleted.” Moreover, as discussed before, there is a structural reason why passive subjects are capable of remaining low in the sentence, because they originate as an internal argument within the VP.

The restriction of intransitive verbs to unaccusatives as opposed to unergatives (cf. Perlmutter 1978) builds on the distinction between unaccusatives and unergatives with respect to the original position of the grammatical subject in the clause—i.e. the position where the argument is base-generated in earlier versions of generative syntax, or merged in Minimalism. While subjects of unergative verbs are taken to be underlying subjects (external arguments), subjects of unaccusative verbs are taken to be underlying objects (internal arguments). Following the observation that unaccusatives but not unergatives allow locative inversion, the possibility to occur in locative inversion is often taken as a diagnostic for unaccusativity (cf. Bresnan & Kanerva 1989, Coopmans 1989, Hoekstra & Mulder 1990, Culicover & Levine 2001). However, there are counterexamples, witness in (8) and (9).

(8) *On the top floor of the skyscraper broke many windows.

(9) On the third floor WORKED two young women called Maryanne Thomson and Ava Brent...[L. Colwin, *Goodbye without Leaving*, 54]

(Levin & Rappaport Hovav 1995: 224)

Example (8) illustrates how not all unaccusatives allow inversion, while (9) provides evidence that some unergatives do allow inversion. Crucially, then, the generalization that inversion occurs with unaccusatives cannot account for all occurrences of locative inversion.

Two further observations about the type of verb are in order. First, some authors make a distinction between inversion with *be*, as in (10), and inversion with other verbs, on the grounds that these are separate phenomena (see Birner & Ward 1998: 183 and Bresnan 1994:76-77).

(10) He was making tea and warming his deeply lined, cracked hands on the pot — *under his ragged nails was the mechanic's permanent, oil-black grime.* [Irving 1985: 291]

(Birner & Ward 1998: 182, their 229)

The most important argument in support of *be*-inversion as a separate phenomenon is that it occurs more freely, allowing inversion after elements that are not PPs and not locative. Second, Birner & Ward (1998) propose that inversion is only possible when verbs are what they call “informationally light” (190), defined as being accessible to the reader and low in information value, as illustrated in (11).

- (11) He opened the door and took a folded canvas bucket from behind the seat.
Coiled on the floor lay a one-hundred-and-fifty-foot length of braided nylon climbing rope three-eighths of an inch thick.

(Birner & Ward 1998: 191, their 241b)

According to Birner & Ward, the verb *lay* is light, because it can be inferred from the context and does not add information—comparable to *be*, which is always light. While Birner & Ward’s account does explain some instances of inversion in which the verb is not unaccusative, it relies rather heavily on interpretation and as such, it is a problematic diagnostic, perhaps in general but most importantly for an historical investigation.

The second aspect of the definition for locative inversion is based on the element in first position: in general, it occurs in sentences where the initial element is a locative PP. As for the type of verb, it is possible to provide counterexamples both for the locative element and for the syntactic category of the initial element. Consider example (12).

- (12) You can drive as fast as you like in the outside lane on a West German highway and may feel like the king of the road — until you look in the rear mirror. *Zooming in on you like a guided missile comes a rival contender, bullying you to get out of the way.* [*Chicago Tribune*, 8/8/89]

(Birner & Ward 1998: 183, their 230a)

Birner & Ward (1998) do not want to use the label locative, but their data show that it is a reasonable generalization: in sentences with inversion and a verb other than *be*, 97% of the initial elements is a locative PP. In addition, while the initial element may not always be a location, strictly speaking, a broader understanding of the term ‘locative’ explains a considerable number of the exceptions, as in Bresnan’s (1994) definition: “[locative] subsume[s] a broad range of spatial locations, paths, or directions, and their extensions to some temporal and abstract locative domains” (1994: 75). In other words, the PP does not need to be a location, but there needs to be at least an indication of location or direction in the sentence. Bresnan uses (13) to illustrate her point.

- (13) a. Crashing through the woods came a wild boar.
 b. *Crashing came a wild boar.
 c. Through the woods came a wild boar.

(Bresnan 1994: 75-76, her 5a and discussion)

Example (13)a does not begin with a locative PP, although the sentence does contain a locative PP. Leaving out this locative PP, as in (13)b, makes the sentence ungrammatical, but leaving out the participle *crashing* does not affect the acceptability, as shown in (13)c. Moreover, Bresnan provides examples which show

that verbs that select locative complements allow inversion while verbs that select no such complements do not allow it, as illustrated in (14)a and (14)b.

(14) a. *Gathered pointlessly in the yard seemed three women.

b. Gathered pointlessly in the yard stood three women.

(Bresnan 1994: 76, her 6a and 7a)

What Bresnan does, then, is expand the notion of location to include sentences in the group of locative inversion even when the locative element is not obvious.

In addition to the initial element and the class of verbs allowing inversion, locative inversion is defined by its discourse or information-structural function. As discussed in Section 2.4.2, Birner & Ward (1998) propose that the function of inversion is to place information that is more given before information that is newer, thereby restoring the unmarked order of information of given-before-new. Crucially, Birner & Ward define this as a relative constraint: there is no absolute condition on the preposed element to be given or the subject to be new.

Other researchers have described the function of locative inversion as ‘presentational’, a function that concentrates on the fact that the subject that follows the verb constitutes new or focused information (Bolinger 1977, Bresnan 1994). Bolinger (1977) describes this function, illustrated in (15), as “present[ing] something on the immediate stage” and “bring[ing] something literally or figuratively BEFORE OUR PRESENCE” (93-94).

(15) Across the street is a grocery.

(Bolinger 1977: 93, his 20)

Bolinger presents this definition in tandem with the definition of another type of sentence, *there*-sentences such as (16), which he analyses as having a comparable presentational function.

(16) Across the street there is a grocery.

(*ibid.* 93, his 21)

The difference between the inversion and *there*-sentences, according to Bolinger, is that the latter “present[...] something to our minds” (94), rather than “before our presence”; in addition, this ‘something’ may already be known in *there*-sentences, while in the inversion sentences the presented element is newly introduced. In contrast to Birner & Ward’s proposal, Bolinger’s account—which Bresnan (1994) also follows—seems to be defined in terms of focus, i.e. the most important information, or the newly connected information (see Lambrecht 1994: 207, 213), instead of information status *per se*.

In conclusion, locative inversion in PDE is generally defined by a sentence-initial locative PP and unaccusative and passive verbs, while there is a range of

counterexamples to this generalization. In some of these cases, the functional character can explain the use of inversion better than the syntactic properties, either the information-rearranging function as proposed by Birner & Ward (1998), or the presentational function as proposed by Bolinger (1977) and followed up by Bresnan (1994). Having determined these properties for PDE, the next question is whether the properties and functions already apply to inversion in earlier periods of English.

5.2.2 Inversion after clause-initial PPs in Early Modern English

This section investigates the loss of inversion after clause-initial PPs in EModE (using the PPCEME, 1500-1710) and MBE (using the PPCMBE, 1700-1914). The first aim of this section is to trace the loss of inversion in PP-initial clauses, in order (i) to compare the loss of inversion in these clauses with the rates of loss of inversion in object and adverb-initial clauses; and (ii) to determine when the PDE pattern of locative inversion emerges. The second aim is to establish whether the influence of information-structural factors in subject placement has changed since the OE period.

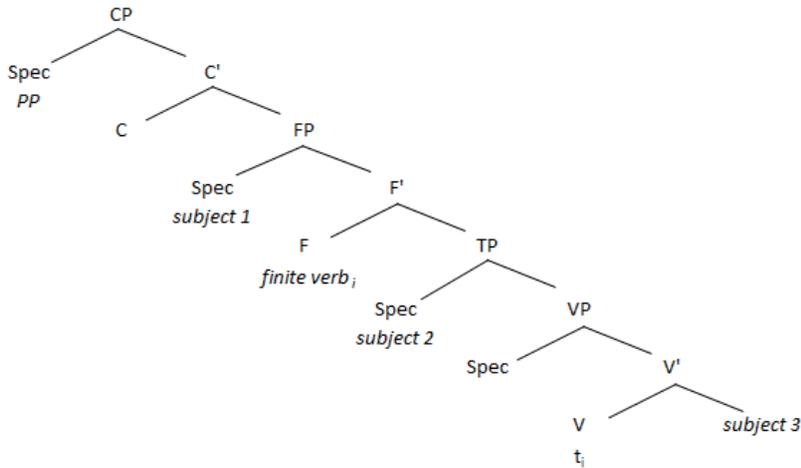
5.2.2.1 Method and selection

Because we are moving from one corpus to another (YCOE to PPCEME), some details in the annotation have changed which need to be incorporated in the selection criteria. In addition, the syntactic changes that have taken place since OE mean that several structural analyses are available for each surface option of inversion or non-inversion, some of which are different from the analyses in OE.

Subject positions

We have already seen (cf. in particular Section 2.2.1 and 4.2.2.4) that OE main clauses which show inversion on the surface may be structurally ambiguous; more specifically, if there are no further diagnostics in the sentence, inverted subjects may be either immediately postverbal (the low subject position) or so-called late (an even lower subject position, remaining within the VP). Figure 5.1 repeats the three subject positions for OE, as identified in the literature.

Figure 5.1 The three subject positions in OE main clauses



Because the PP in Spec,CP is a non-operator, the finite verb moves to F (rather than C), and the subject can occur either before or after it, in Spec,FP or Spec,TP. Representing a separate class of subjects, late subjects originate as complements of the V within the VP and remain in that position.⁶⁷ The word order options for both inversion and non-inversion become more diverse as the Spec,FP subject position is lost over time, and verb movement to F is either lost or replaced with another process, such as V-to-I movement, and the development in the auxiliary and modal system. In addition, some of the diagnostics for the position of either the subject or the finite verb are lost. The following table, inspired by a table from Warner (2007: 90, his Table 5) for fourteenth and fifteenth-century word order patterns, shows the possible analyses of several surface inversion and non-inversion orders after clause-initial PPs in the transition from V2 to the PDE SVO patterns.

⁶⁷ Biberauer & Van Kemenade (2011) point to the possibility of another subject position in the vP domain. They provide the following example where the subject in the subclause is presumably within the vP-domain, but it is not in clause-final position. I will, however, restrict the discussion here to the late subjects in clause-final position in main clauses only.

þa wæs him eallum gesegen, swa swa hit wæs, þæt him wære from Drihtne sylfum heofonlic gifu forgifen

then was them all seemed so as it was that him were by the-Lord himself heavenly grace granted
 'Then it seemed to them all, as indeed it was, that a heavenly grace had been vouchsafed him by the Lord himself.' (2011: 24, their 12c)

Table 5.1 Possible analyses of inversion and non-inversion surface orders after PPs

	Topic position	High position for finite verb	High subject position	Low position for finite verb	Low subject position	PDE position for finite verb	Late subject position	Surface inversion
	Spec,CP	C	Spec,FP	F	Spec,TP	T	VP-complement	
1	PP			FinV	Sbj			Y
2	PP			FinV			Subject	Y
3	PP					FinV	Subject	Y
4	PP		Sbj	FinV				N
5	PP		Sbj			FinV		N
6	PP				Sbj	FinV		N

Let us first consider the three possible analyses for a sentence showing inversion of subject and finite verb on the surface, i.e. those sentences in which the subject appears after the finite verb, given in rows 1 to 3 in the table. First, row 1 illustrates the OE possibility of inversion after non-operators: the finite verb is in F and the subject occurs in Spec,TP, as illustrated in (17) and (18). As discussed in Chapters 2 and 4, this is the word order option in OE which is to a large extent triggered by the information status of the subject (see Van Kemenade 2012), the type which Hinterhölzl & Petrova (2010) call “accidental V2”. In this and the following examples, the initial PP is given in bold, the subject is underlined and the verb is underlined with a bold line.

- (17) **On *bæm gewinne*** *ofslog* *Antipater* *his modor, Cassandres lafe*
 in the strife killed Antipater his mother, Cassander’s widow
 ‘In the strife Antipater killed his mother, widow of Cassander’
 (coorosiu,Or_3:11.82.1.1634)

- (18) **On *ðæm dagum*** *wilnade* *sum æðeling* *to ricsianne in Argentine bære ðeode,*
Falores wæs haten.
 in those days wished some prince to rule in Agrigentum the country Phalaris
 was called
 ‘In those days, a certain prince called Phalaris, wished to rule in the country
 of Agrigentum.’
 (coorosiu,Or_1:12.33.34.657)

In (17), the subject *Antipater* occurs after the finite verb *ofslog*, while its occurrence before the VP-internal object *his modor, Cassandres lafe* works as a diagnostic to determine that the subject has been raised out of the VP. Example (18) shows a

similar word order and structure, but with a different diagnostic: the subject *sum æðeling* occurs after the finite verb *wilnade* but before the VP-internal non-finite verb *to ricsianne*.

The two other possible analyses for surface inversion are late subject orders, with the subject occurring in its base position as complement of VP and the verb either in F or T, as illustrated in rows 2 and 3, respectively. There are certain diagnostics available to determine the position of the subject as late, rather than immediately postverbal: if the subject follows VP-internal material, such as the non-finite verb, it is late, i.e. the subject itself has remained within the VP. However, no such straightforward diagnostics are available for the position of the verb as being either in F or T. This means that there remain sentences that are structurally ambiguous between the structures shown in rows 1-3, as in (19)-(21).

- (19) *On bæm morum eardiað Finnas.*
 ‘In these mountains [wastes] dwell Finns’
 (coorosiu,Or_1:1.15.24.269)
- (20) *On the side of Were river is Stanop.*
 (leland-e1-h,70.31)⁶⁸
- (21) *At her reward dined my cousen Fraunces and my cousen Maqret.*
 (edward-e1-h,363.285)

The sentences in (19)-(21) lack the diagnostic elements and it is not possible to tell whether the subject is late (row 2 or 3), or immediately postverbal (row 1). However, as V2 is lost, i.e. verb movement to F is lost and the subject position in Spec,FP is either lost or collapses with Spec,TP (cf. Hinterhölzl & Petrova 2010; Van Kemenade 2012), the analysis in row 2, with the finite verb in T and the subject within the VP, becomes the only analysis that remains available for surface inversion. Note that subject-auxiliary inversion—as in (2)—is a separate phenomenon, in which V continues to move to C right through to PDE (i.e. in I-to-C movement), and is hence not affected by the loss of FP.

The surface orders which do not show inversion of the subject and finite verb occur in any of the structures illustrated in row 4-6. The first option, illustrated in row 4, represents the traditional V3 environment in the OE analysis.

- (22) *On ðæm gefeohte Pirrus hæfde eahtatig M feþena, & V M gehorsedra.*
 ‘In that battle Pyrrhus had eighty thousand feet, and five thousand horses’
 (coorosiu,Or_4:1.86.1.1731)

⁶⁸ In the examples from the PPCEME, the subperiod is given in the file name, with E1 representing the years 1500-1569, E2 1570-1639 and E3 1640-1710.

In (22), the subject *Pirrus* occurs in Spec,FP, before the finite verb *hæfde* in F. This is the standard analysis of these orders as long as V2 is in place, i.e. as long as the verb moves to F in non-operator contexts. Row 5 shows an analysis which should be logically possible as V2 is lost, but it is an unlikely one. In this analysis, the subject occurs in the same position as in the previous analysis, Spec,FP, but the finite verb occurs lower than in the previous analysis, in T rather than in F. However, if verb movement in V2 contexts is taken to be movement to a position which subsequently needs its specifier to be filled (cf. Hinterhölzl & Petrova 2010), this analysis is out because head and specifier of the same projection are not filled in row 5. Interestingly, Van Kemenade (2012: 833) in fact proposes that the position of the verb may be less important than the position of the subject in accounting for the loss of V2: rather than the loss of verb movement, the loss of V2 is the result of a change in the position of the subject, in the absence of evidence for a change in the position of the finite verb. The final loss of inversion as far as it concerns verb movement takes place only after the ME, and Van Kemenade hypothesizes that this coincides with the loss of V-to-I movement.

Finally, the order illustrated in row 6 is the one that increasingly becomes the standard order when movement of the finite verb to F is lost, with the finite verb in T and the subject in Spec,TP. In fact, it seems the entire projection F is lost—object pronouns and subjects no longer occur in the specifier, which then also leaves little evidence for the verb in F (although there is some evidence from negation contexts, see Fischer et al. 2000: 135; Rissanen 1994, 1999). In conclusion, the diachronic development over the course of the centuries consists of the loss of two operations that characterized the OE system: loss of verb movement to F and loss of the subject position in Spec,FP, possibly because it collapses with Spec,TP.

Queries and selection

In order to ensure that the data presented in this chapter can be compared to those in Chapter 4, the selection of clauses is similar to the previous corpus studies. This means that, following the observation that V2 is a main clause phenomenon⁶⁹ and that the variation in subject positions only applies to nominal subjects, only main clauses with nominal subjects were selected. The selection is different in one important respect: while in OE, a distinction was made between conjunct and non-conjunct clauses because of the consensus in the literature that conjunct clauses disfavour inversion and the results for the study of PP-initial clauses in OE (Section 4.4.3, but see also the discussion in Section 4.3.4 for object-initial clauses), this distinction seems to have become less of a determining factor once we move to EModE. I will therefore only comment of the difference between conjunct and non-

⁶⁹ It is not just that V2 is a main clause phenomenon. Biber et al. (1999: 926) find that inversion is a main clause phenomenon in their PDE corpus: over 90% of the inversions in their corpus of conversation, fiction and news occur in main clauses. For academic prose, this is 75%.

conjunct clauses if relevant. Query 5.1 presents the basic query that was used; the detailed selection of finite verbs will be explained below.

<i>Query 5.1</i> Basic query for inversion and non-inversion in PP-initial main clauses
a. IP-MAT idoms finite verb <i>and</i> IP-MAT idoms NP-SBJ* <i>and</i> IP-MAT idomsfirst PP*
b. PP* iprecedes finite verb <i>or</i> NP-SBJ*
c. Inversion: finite verb precedes NP-SBJ*
d. Non-inversion: NP-SBJ* precedes finite verb

Line 5.1a selects main clauses with a finite verb, a subject, and a prepositional phrase in clause-initial position. Line 5.1b states that the PP should immediately precede the finite verb or the subject, excluding clauses where the PP is followed by for instance an adverb (cf. Section 4.4.4). Lines 5.1c-d present the queries that were used to distinguish between inversion and non-inversion of the finite verb and the subject. Note that the query for inversion does not distinguish between immediately postverbal (V2-type of inversion) and late subjects.

There are three aspects of the selection of clauses that are different from the selection in Chapter 4. First of all, not only do we need to exclude pronominal and ellipted subjects, but we now also need to exclude expletive *there* subjects, as they almost invariably precede the finite verb.⁷⁰ The following query was used to select pronominal and expletive subjects, which were then excluded from the database.

<i>Query 5.2</i> Selection of non-nominal subjects
a. IP-MAT idoms NP-SBJ* <i>and</i> NP-SBJ* idomsonly *PRO* <i>or</i> *EX*

The complement file this query, i.e. all sentences that did not contain a pronominal or expletive subject, was used as input for further queries.

A further distinction was made for type of finite verb, which was possible because of the availability of a corpus annotated for unaccusativity.⁷¹ The following classes of verbs were distinguished: finite auxiliaries, unaccusative finite verbs and a ‘rest’ category. Following Van Kemenade (2012), I will call this last category transitive/unergative, but it should be noted that this group may inadvertently include a few unaccusatives that were not detected, for instance because of extremely unusual spellings. The group of unaccusatives includes impersonals, weather verbs, raising verbs, “verbs of coming and going”, inchoatives, posture verbs, manner of motion verbs (Van Kemenade 2012: 828). Query 5.3 shows the

⁷⁰ Only in E1 are there a number of sentences in which *there* occurs after the finite verb (10 examples, 8.3% of the total 121 examples in that period). No examples with inverted *there* occur in E2 and E3 in this dataset.

⁷¹ This corpus was kindly made available to me by Ans van Kemenade (for details, see Van Kemenade 2012; Van Kemenade & Westergaard 2012).

definitions for each of the verb categories, and the order in which the queries to select these groups of verbs were carried out.

Query 5.3 Selection of different types of finite verbs	
a. On source file:	IP-MAT idoms finite verb
b. On output line 1:	IP-MAT idoms unaccusative verb
c. On complement file line 2:	IP-MAT idoms finiteaux
d. On complement file line 3:	IP-MAT idoms finite verb
<p><i>Definitions:</i></p> <p>finiteverb: BEI BEP* BED* UTP *HVI *HVP* *HVD* *AXI *AXP* *AXD* *MD* VBI *VBP * *VBD* *DOI *DOP* *DOD* NEG+BEI NEG+BEP* NEG+BED* NEG+AXI NEG+ *AXP* NEG+*AXD* NEG+*MD NEG+VBI NEG+*VBP* NEG+*VBD</p> <p>finiteaux: BEI BEP* BED* UTP *HVI *HVP* *HVD* *AXI *AXP* *AXD* *MD* *DOI *D OP* *DOD* NEG+BEI NEG+BEP* NEG+BED* NEG+AXI NEG+*AXP* NEG+*AXD * NEG+*MD*</p> <p>ufiniteverb: BEI BEP* BED* U-BEI U-BEP* U-BED* U-VBI* U-VBP* U- VBD* NEG+BEI NEG+BEP* NEG+BED*</p>	

Line 5.3a first selects all clauses with a finite verb, while lines 5.3b-d select the different types of finite verbs: 5.3b selects unaccusative verbs, 5.3c auxiliaries and 5.3d all remaining verbs. In order to ensure that no verbs were selected twice, the complement files were used in this series of queries. Because *be* is included as an unaccusative, i.e. selected in line 5.3b, passives are grouped with unaccusatives.

Finally, several queries were added to clean up the category of clause-initial PPs. First, there are two different ways in which PDE referential adverbs like *thereby* and *therefore* are coded as PPs, possibly reflecting an intermediate state in their lexicalization, where their origin of adverb and P combination is still clearer than it is today.

(23) [IP-MAT [PP [ADV *There*] [P *by*]] *ys the place and a stone lying wher our blyssyd lady
 Died and assumptyd In to hevyn.*
 (torkingt-e1-h,35.186)

(24) [IP-MAT [PP [ADV+P *Therefore*]] *power, is desired,*
 (boethco-e1-h,76.373)

In (23), *There* is coded as an adverb, combining with the preposition *by* to make a PP. In (24), *Therefore* is coded as one word, a combination of adverb and preposition,

but this combination is coded as a PP rather than an ADVP. I selected these PPs with the following query, and excluded them from the database.

Query 5.4 Adverbs coded as PPs
a. IP-MAT idomsfirst PP*
b. PP* idoms ADV+P ADV

These adverbs-coded-as-PPs are found reasonably frequently in E1 and E2, representing 3.9% and 4.2% of the total number of PP-initial clauses with a finite verb and a nominal subject (91/877 in E1 and 126/854 in E2), but their number is considerably lower in E3, with only 4.7% (29/621), presumably because they are increasingly coded as adverbs instead of PPs.

Another type of PP also needed to be excluded because, as in the YCOE, several subordinate clauses are coded as PPs in the corpus, illustrated in the following example.

- (25) [_{IP-MAT} [_{PP} [_P *when*] [_{CP-ADV} [_C 0 [_{IP-SUB} *blyssednes semeth to contayne many thyngs*,]]]]] *it is to doubtte, whether all these thynges do ioynе togyther, ...*]
 (boethco-e1-h,76.311)

In (25), the subordinator *when* is coded as a P, while its complement is a clause, CP-ADV with an empty (“0”) head. These clauses were selected with the following query and excluded from the database.

Query 5.5 Clauses coded as PPs
a. IP-MAT idomsfirst PP*
b. PP* idoms CP*

These clauses are in fact quite common in the PPCEME, making up between 22.6% and 30.0% of all PP-initial clauses with a finite verb and a nominal subject (211, 193 and 186 examples respectively), so failure to exclude these would lead to a serious problem with the quality of the database of PP-initial clauses.

5.2.2.2 The loss of inversion after PPs

This section presents a study of the rates of inversion in PP-initial clauses in the EModE corpus, which begins at 1500, the time that is usually taken to present the final stages of the loss of V2 (see Fischer et al. 2000, Haerberli 2002). However, at the same time there is variation between contexts, most notably depending on the type of initial element and the type of verb, as shown by Van Kemenade & Westergaard (2012). The aim of this section is to establish at what speed inversion after clause-initial PPs is lost in the EModE period, and crucially, in view of the PDE pattern of

locative inversion, whether there is evidence for the beginnings of such a pattern, i.e. whether there are specific contexts in which inversion is lost at a lower speed.

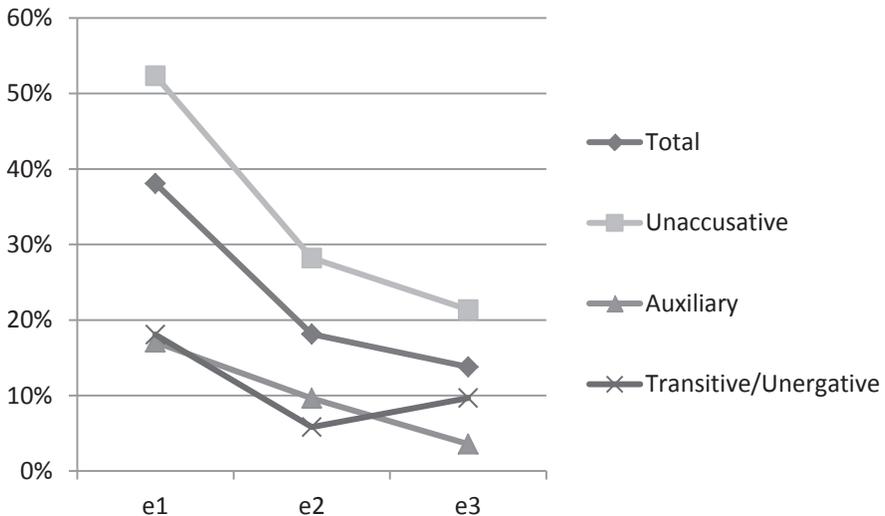
Table 5.2 and Figure 5.2 present the results for the EModE period, based on the coding for unaccusative verbs, auxiliary verbs and the transitive/unergative group as outlined in the previous section.

Table 5.2 Rates of inversion in PP-initial clauses in the PPCEME according to verb type

	E1			E2			E3		
	<i>n</i> inv	<i>n</i> total	% inv	<i>n</i> inv	<i>n</i> total	% inv	<i>n</i> inv	<i>n</i> total	% inv
Unaccusative	177	338	52.4	79	280	28.2	43	201	21.4
Auxiliary	14	82	17.1	8	83	9.6	4	112	3.6
Transitive/ unergative	28	155	18.1	10	172	5.8	9	93	9.7
Total	219	575	38.1	97	535	18.1	56	406	13.8

E1: $\chi^2(2) = 35.68$, $p < .001$; E2: $\chi^2(2) = 40.79$, $p = 0$; E3: $\chi^2(2) = 20.93$, $p < .001$

Figure 5.2 Rates of inversion in PP-initial clauses in the PPCEME according to verb type



The line with the diamonds in Figure 5.2 shows the percentage of inversion for all types of verbs, starting with 38.1% in E1, still quite a high percentage considering this is the time that V2 is generally said to be lost already. This line shows a quite

steep decline from E1 to E2 (from 38.1% to 18.1%). This decline is echoed in the results for the different types of verbs, where it is most pronounced for the unaccusative verbs: from 52.4% in E1 to 28.2% in E2. After this initial decrease, the percentages for unaccusative and auxiliary verbs continue to decline. The category of transitive/unergative verbs, however, increases from 5.8% in E2 to 9.7% in E3. While there might truly be an increase which requires an explanation, it should be noted that, since this is the category of uncoded or 'rest' examples, the category contains mixed groups of verbs, and there are at least some examples that turn out to be unaccusatives, such as (26). This group also contains at least one example with *said*, given in (27), another verb known to favour inversion.

(26) *In the meanetyme commithe sir Edwarde Hastynqes, newly made master off the horse to the queene,*
(underhill-e1-p2,136.25)

(27) *In the second of Haggiae, speake now, saith God to his prophet,*
(hooker-b-e2-P2,55.85)

Interestingly, this increase in the percentages of transitive/ergative examples is stronger in non-conjunct clauses (4.9% to 15.0%) than in conjunct clauses, which show a decrease (7.2% in E2 to 0.0% in E3). The same effect is seen in the unaccusative verbs: they show a slight increase from E2 to E3 in non-conjunct clauses (22.8% to 24.1%) but a further decrease in conjunct clauses (38.0% to 16.2%). These two patterns represent the only two differences between conjunct and non-conjunct clauses in this database, but the data presented here are too limited to provide any real insight into the difference between conjunct and non-conjunct clauses.

From these data, the effect of the type of verb is clear, with unaccusatives generally showing a higher percentage of inversion throughout the EModE period. At the same time, these unaccusatives are also responsible for the main drop from E1 to E2, showing a drop from 52.4% to 28.2%. There are several types of clauses among the unaccusative examples in E1 that might go some way towards an explanation of this substantial decrease from E1 to E2. First, there is one clear pattern that stands out, illustrated in (28) and (29).

(28) *To the worshipfull Robart Plompton, knight, be thes byll delivered in hast.*
(aplumt-e1-h,171.52)

(29) *To Sir Richard Plompton be thes byl delivered in hast.*
(aplumt-e1-p1,187.12)

These sentences, with a preposition plus a name, followed by the passive *be delivered*, seem to represent a fixed construction, supported by the fact that this is a

mode of address found in several letters in the corpus. Because of their fixed structure, they can easily be selected with a query, given below.

<i>Query 5.7 Selection of be this delivered sentences</i>
a. IP-MAT idoms NP-SBJ* <i>and</i> NP-SBJ* idoms D*
b. IP-MAT idoms *VAN* <i>and</i> *VAN* idoms delivered

In this query, line 5.7a selects demonstrative subjects, while line 5.7b—executed on the output of line 5.7a—selects those clauses with a passive participle *delivered* (I found no other spelling variants). The query was executed on the entire PPCEME. Note that this query selects both clauses in which the demonstrative subject follows the participle and clauses in which it precedes the participle.

There are 31 sentences in total in the entire PPCEME with *be this (letter/bill) delivered* or the non-inverted version *this (letter/bill) be delivered*, crucially all but one occurring in the E1 period. The preference for the inverted version is clear: 23 examples show inversion against 8 with no inversion. They all occur in the unaccusative group in Table 5.2 because the finite verb in all these sentences is *be*. If this specific pattern is excluded from the results, the overall percentage of inversion with unaccusative verbs becomes slightly lower in E1: 48.9% instead of 52.4% (154 of 315 examples). Although it affects the percentage, the pattern does not by itself account for the high percentage of inversion with unaccusative verbs, with 48.9% still a considerably higher percentage than shown for the other groups of verbs.

Another observation with regard to the E1 unaccusative examples with inversion is that they include some types of clauses that are no longer grammatical in PDE, again indicating that there may be a particular type of sentence which leads to the high percentage of inversion in E1 and whose loss would explain the decrease throughout the rest of the EModE period. The most important type in this group consists of clauses that show the old V2 type of inversion, where the finite verb has moved in front of the subject but the participle remains low, within the VP, as in (30)-0. The subject in these examples is not late, but presumably in Spec,TP, the V2 position for inverted subjects.

(30) *to this folye were theyse frerys brought, by the excytinge of the duke of Burgoyne, as the common fame went.*

(fabyan-e1-p1,558.51)

(31) *In this yere also was dame Isabell, some tyme wyfe of Rycharde, lately kynge of Englonde, married vnto Charlys, eldest sone of the duke of Orleau~ce;*

(fabyan-e1-p1,559.72)

- (32) *In this yere, and moneth of lanuary, were certayn courses of warre ron in Smythfelde atwene sir Edmu~de erle of Kent, and the lorde Moryf a baron of Scotlande, vpon the chalenge of the sayd Scottisshe lorde;*
(fabyan-e1-p2,572.44)

Inversion in these sentences may be used for a discourse effect: Los (2013) finds that V2 is often used as an episode boundary marker, especially after phrases such as *in this year* in Capgrave Chronicles, a fifteenth-century text. Excluding the *be this delivered* sentences, which have the same structure, there are 37 examples with the old V2-type inversion in E1, 11 in E2, and 2 in E3, which explains part of the decline of inversion after unaccusative verbs from E1 to E2.

Two other types of sentences with unaccusative verbs—subjunctives and late subjects—become less frequent over the EModE period, although in contrast to the examples shown in (30)-0, these remain possible in PDE to some extent. The first of these, the subjunctives, are illustrated in (33), and are also used in the *be this delivered* and the *be it enacted* types.

- (33) *In the paroch chirch of S. Brandon at Branspeth be dyvers tumbes of the Nevilles.* (Ieland-e1-h-H,72.65)

The following table shows the frequency of these subjunctives over the course of the EModE period.⁷²

Table 5.3 Frequency of subjunctives in PP-initial clauses with unaccusative verbs

	E1	E2	E3
No. of subjunctives	48	6	2
Total no. of PP-initial clauses with an unaccusative verbs	338	280	201
% of subjunctives	14.2	2.1	1.0

$\chi^2(2) = 49.23, p = 0$

In E1, where the overall frequency of subjunctives is still quite considerable (14.2%), the majority of subjunctives occur in an inversion context (60.4%, 29 of the 48 examples). The decrease in these subjunctives in PP-initial clauses reflects an overall development of a drop in number and proportion of subjunctives, with a decline from 1.9% in E1 (522/28202) to 0.4% in E3 (96/24970) in all main clauses.⁷³ The second type of clause that becomes less frequent are the late subjects in passive sentences, as illustrated in (34) and (35).

⁷² Subjunctives are not coded separately in the PPCEME, so I used the following queries, selecting all finite forms of *to be* where the actual form of the finite verb was *be*: IP-MAT idoms BE|BEP*|BED*, and BE|BEP*|BED* idoms be|Be. The query was executed on PP-initial clauses with unaccusative verbs.

⁷³ The same query was used as given in fn.71, but now on all clauses and compared to the total number of main clauses (IP-MAT) calculated by *CorpusStudio*.

- (34) *with this most precious blud was illumined the fyue great Capital letters in this wonderful booke.*
(fischer-e1-p1,396.309)
- (35) *about the corps were borne four banners, a banner of the order, another of the red rose, another of queen Jane Seymour, another of the queen's mother.*
(machyn-e1-h,40.48)

The following table shows how, within the category of unaccusative verbs in PP-initial clauses, passive main clauses with late subjects become less frequent over the EModE period, starting at 11.8% in E1 and declining to 4.0% in E3.

Table 5.4 Frequency of late subjects in PP-initial clauses with passive verbs in PPCEME

	E1	E2	E3
Number of late subjects in passive clauses	40	23	8
Total number of clauses with PP and unaccusative verbs	338	280	201
% of late subjects	11.8	8.2	4.0

$\chi^2(2) = 9.93, p < .01$

Although this table shows a decrease in the frequency of these type of passive, we have already seen in Section 5.2.1 that they still occur in PDE. What these specific categories of examples and their decrease in frequency show is that even though unaccusatives continue to show inversion at a higher rate, specific subtypes (subjunctives, late subjects of passive clauses) are lost or become less frequent over the course of the EModE period, thus accounting for the decrease in percentages in the unaccusative category. In other words, while inversion with unaccusative verbs remains possible, it becomes more limited over time.

The group in Table 5.2 with the lowest percentages of inversion throughout comprises transitive/unergative verbs; indeed, this is the group that generally does not occur with inversion in PDE. That being said, the few examples that do occur with inversion provide additional insight into the principles of inversion in this period. The first observation about these examples is the weight of the subject: a number of these sentences have a heavy subject, in turn an important consideration for having a subject late in the sentence.

- (36) *At the gate therof met her the lady marcus of Northampton, the countesse of Penbroke, and divers other ladies and gentlewomen to the number of threscore.*
(edward-e1-h,359.244)

- (37) *Vpon [the] v. day played togyder an Henauder and a Squyre called Iohn~ Steward*
(fabyan-e1-h,173V.C1.180)
- (38) *In the selfe same day entred Noah, and Sem, and Ham, and Japheth, the sonnes of Noah, and Noahs wife, and the three wiues of his sonnes with them, into the Arke*
(authold-e2-h,7,1G.368)

In all three sentences, the subject is a list of names, making it long and complex in structure. This means that it is likely that the motivation for the late occurrence of the subject in these sentences lies with principles of end weight, as in for example Heavy NP shift (see Wasow (2002), who follows Quirk et al. (1972) and Hawkins (1994)).

The second observation is that in at least some of the examples, the text seems to be a determining factor, in that it is only certain texts that show inversion with transitive/unergative verbs while others do not. For example, there are 5 inversion examples in E3 that occur in the same text, and moreover have the same verb, *preach*. These examples make up over half of the 9 examples in E3.

- (39) *At St. James's church preached Dr. Burnet, on 5. Deut:*
(evelyn-e3-h,901.125)
- (40) *At St. Martines preached Dr. Tenison on: 2. Cor: 4. 8.*
(evelyn-e3-h,904.183)
- (41) *In the Afternoon preached the B: of Banqor: on 37: Psal: 37:*
(evelyn-e3-p1,916.311)
- (42) *In the afternoone preached the Deane of St. Asaphs on: 1. John: 5. 4:*
(evelyn-e3-p1,922.446)
- (43) *In the afternoone Preached for Dr. Tenison, at St. James's new church Dr. Bohune the same sermon he preached at Deptford, the 10th of the last moneth:*
(evelyn-e3-p2,884.164)

Not only do these examples, with the verb *preach* followed by the name of a preacher and the content of the sermon, work as a kind of fixed format within this text, they also seem to be typical of the presentative function as defined by Bolinger (1977): they bring the name of the preacher to the stage (or the pulpit), and possibly also the topic of the sermon. The fact that examples such as these are limited to

certain texts means that the occurrence of inversion with transitive/unergative verbs cannot be generalized to the entire language at this stage.

Finally, there are some examples that build on the OE function of local anchoring in the clause-initial position, in turn prefiguring the PDE pattern of locative inversion occurring specifically after anaphoric links (cf. Biber et al. 1999: 911, 914).

(44) *to this belonged Three stately Tanks with a deep Well, some an Hundred Fathoms down to the Bottom, with a Pond, either to Water or Wash the Cattel in, hard by it; the perfect proof of our being passed the Sands, which hold not digging such a depth, this being White Marle;*

(fryer-e3-p1,2,198.192)

(45) *To this place belong two sorts of Vermin, the Fleas and Banyans;*

(fryer-e3-p2,1,211.47)

The clause-initial PPs provide a starting point in first position and a new element is introduced after the verb, corresponding to the presentational function that was proposed by Bolinger (1977) and Bresnan (1994) for PDE, and also to Birner & Ward's (1998) relative information-ordering function for inversion. In these cases the order of information, or more specifically the presentation of new information may override the syntactic principles of inversion.

In conclusion, inversion after all types of verbs becomes decidedly less frequent throughout the EModE period, with the overall percentage going down from 38.1% in E1 to 13.8% in E3. There are considerable differences between the groups of verbs, with unaccusatives showing the highest percentages throughout and remaining at 21.4% in E3. At the same time, these unaccusatives are responsible for the largest drop in the percentages from E1 to E2, from 52.4% to 28.2%. I have explained this drop from E1 to E2 by identifying specific types of patterns within the group of unaccusatives that become less frequent throughout the EModE period, most notably the old V2 type of inversion, where the subject occurs after the finite verb but before the non-finite verb. What has become clear is that these old patterns are lost, while at the same time some aspects of the PDE pattern of locative inversion are visible in the remaining examples: the high frequency of unaccusative verbs in inversion contexts, as well as the presentative and information-structural function of inversion. This pattern comes to stand out more as the remnants of the old V2 type of inversion are lost over time.

5.2.2.3 Information-structural factors in subject placement in E1

The syntactic changes that have taken place involving the position of the subject and finite verb also have consequences for the influence of information-structural factors on subject placement. Most importantly, the variation that existed in OE between

the two subject positions in the variable V2 context is starting to disappear in the EModE period. From an information-structural perspective, the fact that the syntactic variation between preverbal (Spec,FP) and immediately postverbal (Spec,TP) disappears means that there is no longer a role to play in this area by information-structural factors: Spec,TP is developing into the default subject position, regardless of the information status of the subject. What does remain similar to OE, however, is that there is syntactic variation between subjects in Spec,TP and late subjects, with an information-structural motivation for the late subjects.

This study is based on a selection of the PP-initial main clauses with nominal subjects that were used in the study in the previous section. In order to allow for a qualitative analysis of the examples, the selection of texts was limited to the Helsinki part of the PPCEME, the E1 subperiod, and non-conjunct clauses (also, in this period, to avoid any possible influence of conjunctions), resulting in a dataset of 100 PP-initial main clauses with nominal subjects. As in the previous sections, clause-initial subordinate clauses and adverbs coded as PPs in the corpus were excluded, as were ellipsed, pronominal and expletive subjects. The clause-initial PP should immediately precede either the finite verb (in an inversion context) or the subject (in a non-inversion context). I have not initially, for this study, distinguished between different types of verbs or the different positions of the subject that occurs after the finite verb, i.e. between immediately postverbal and late subjects, although these factors will be considered in the discussion of the examples. For the analysis of the information status of the subject the same categories as in the studies in Chapter 4 were used. Table 5.5 repeats Table 4.2, giving an overview of the information status categories.

Table 5.5 (=Table 4.2) Annotation scheme for information status

Category	Categories from other models included in each category	Binary distinction
OLD	Prince's (1981) <i>Textually Evoked</i> , Komen et al.'s (2014) <i>Identity</i> , Pintzuk & Taylor's (2011, 2014) <i>Previously mentioned</i> .	'Given'
ACCESSIBLE	Prince's <i>Unused</i> and <i>Situationally Evoked</i> , Komen et al.'s <i>Assumed</i> , Taylor & Pintzuk's <i>Shared/Cultural knowledge</i> , <i>Situationally Evoked</i> , Haug et al.'s (2014) <i>Generic</i> , Birner's (2006) <i>Elaborating inferables</i>	
ANCHORED	Birner's <i>Bridging inferables</i> , Prince's <i>Brand-new Anchored</i>	'New'
NEW	Prince's <i>Brand-New Unanchored</i> , Taylor & Pintzuk's and Haug et al.'s <i>Specific</i> , <i>Non-specific</i> and <i>Short-term referents</i> , Komen et al.'s <i>Inert</i>	

Table 5.6 shows the results of the analysis of the information status of preverbal and postverbal (both immediately postverbal and late) subjects in main clauses with all types of verbs, presenting not only the results for each information status category, but also the binary division between ‘given’ (OLD and ACCESSIBLE) and ‘new’ (ANCHORED and NEW).

Table 5.6 Information status of pre- and postverbal nominal subjects in PP-initial clauses in E1

	Non-inversion		Inversion		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
OLD	28	70.0	12	30.0	40	100.0
ACCESSIBLE	16	64.0	9	36.0	25	100.0
Subtotal ‘Given’	44	67.7	21	32.3	65	100.0
ANCHORED	3	60.0	2	40.0	5	100.0
NEW	7	23.3	23	76.7	30	100.0
Subtotal ‘New’	10	28.6	25	71.4	35	100.0
Total	54	54.0	46	46.0	100	100.0

$\chi^2(3) = 16.56, p < .001$; between ‘Given’ and ‘New’: $\chi^2(1) = 14.02, p < .001$

The table shows that the majority of OLD or ACCESSIBLE subjects occur before the verb (67.7%) and that the majority of ANCHORED and NEW subjects occur after the verb (71.4%). These percentages are comparable to the percentages in Section 4.4.4 for the subjects in PP-initial clauses in *Orosius*, with roughly two-thirds of ‘given’ subjects inverted and the same percentage of ‘new’ subjects showing uninverted order. As in OE, this supports a general tendency for given subjects to occur early and new subjects to occur late, although there is no direct relation between information status and position. Considering the number of counterexamples in Table 5.6, it may seem that the tendency for given-before-new is not strong, but in fact many of these counterexamples are independently motivated and represent cases where information structure is overruled by formulaic conventions and deliberate archaisms.

The first unexpected pattern from the perspective that given subjects occur before the finite verb and new subjects follow it is represented by subjects which are ANCHORED or NEW and which occur before the finite verb. There are 10 such examples in this database. Half of these are best analysed as event-reporting sentences (cf. Lambrecht 1994: 124): these sentences do not have a regular topic-comment or focus-presupposition structure, but present new information entirely, often as a discourse-initial sentence. Two of these are given in (46) and (47).

- (46) *In this yere and moneth of Noue^{bre}bre one named the Walshe Clerke apeched a knyght called Sir Percyvall Sowdan of Treason for tryall*
(fabyan-e1-h,172V.C2.143)

- (47) *Upon a tyme certayn women in the countrey were appoynted to deryde and mokke a frere a lymytour that vsyd moche to vysyth them.*
(merrytal-e1-h,135.357)

In (46), both the subject *one named the Walshe Clerke* and the object *a knight called Sir Percyvall Sowdan* are new to the discourse, and the sentence is the first of a new section after a section break. In (47), again, the subject *certayn women in the countrey* is new, as is the event related in the sentence. Like the first example, it occurs as the first sentence of a new section, a new ‘merry tale’. The clause-initial PPs in these two examples are frame setters: they provide the context for the following statement and although there are no contrastive settings in the text, there is an implicit contrast as the frame setter limits the context to this specific one (cf. Krifka 2007). The event-reporting nature of these sentences explains the occurrence of a new subject in a preverbal position because the consideration that given precedes new does not play a role in a sentence that presents only new information. Not all of the preverbal new subjects occur in event-reporting sentences, however. One possible explanation for their preverbal occurrence is that these are the first signs of a default subject position beginning to emerge, considering that overall less than half of the clauses show inversion (46.0%, see Table 5.5).

The second unexpected pattern is represented by ‘old’ subjects occurring after the verb, 21 examples in this database. These sentences are structurally ambiguous between late subjects (within the VP) and the V2-type of inversion. Inversion in these cases can, at least to some extent, be explained with reference to the factors identified in the previous section as promoting inversion. First, almost half of the ‘given’ inverted subjects (10 out of 21) are *To x, be this delivered* sentences, representing a specific and fixed pattern.

- (48) *To the worshipful Sir Robart Plompton, kt. Be thes delivered in hast.*
(aplumpt-e1-h,167.3)

The question for these examples is to what extent the information status is relevant: not only is this a fixed formula, which may rule out any information-structural effects, but there is also a deictic element present, with a reference to the real-world situation, both for the subject *thes* and the clause-initial PP *To the worshipful Sir Robert Plompton, kt.*, rather than the world of the text, which means that the information status within the text is difficult to determine. The second explanation for the occurrence of old subjects after the finite verb comes from the type of verb: 17 of the 21 examples (including the *to x be this delivered* sentences) are unaccusatives.

Interestingly, three of the four examples in this database that are OLD or ACCESSIBLE and occur after the finite verb but do not contain an unaccusative verb are examples of the old V2 type of inversion: they occur after the finite verb but before

other VP-internal material. In other words, for these sentences it is possible to determine the position of the subject as immediately postverbal rather than late.

- (49) *To this bath do gentilmen resort.*
(Ieland-e1-h,142.302)
- (50) *from that place hath the spirite of feeling his first creation,*
(vicary-e1-h,32.62)
- (51) *By them two, must the Kynge se euer when he hath to muche.*
(latimer-e1-h,38L.355)

In (49), the subject *gentilmen* occurs after the finite verb *do* but before the non-finite verb *resort*. In (50), the subject *the spirite of feeling* occurs after the finite verb *hath* but before the object *his first creation*. And finally, in (51), the subject *the Kynge* occurs after the finite verb *must* but before the non-finite verb *se* ('see'). In all three sentences, then, the subject precedes VP-internal material and must therefore be in Spec,TP with the finite verb in F.

Finally, this selection of examples sheds a light on the emergence of a pattern of a presentational construction with clause-initial PPs. Some examples in the database clearly introduce a NEW referent in the late subject position, as in (52) and (53).

- (52) =(36)
At the gate therof met her the lady marcus of Northamton, the countesse of Penbroke, and divers other ladies and gentlewomen to the number of threscore,
(edward-e1-h,359.244)
- (53) *In the paroch chirch of S. Brandon at Branspeth be dyvers tumbes of the Nevilles.*
(Ieland-e1-h,72.65)

The subject of (52), *the lady marcus of Northamton, the countesse of Penbroke, and divers other ladies and gentlewomen to the number of threscore*, is NEW, as is the subject of (53), *dyvers tumbes of the Nevilles*. While the persons mentioned in example (52) are not immediately continued in the text, the text from which (53) is taken goes on to elaborate on these *dyvers tumbes* and explains which people lie there. These, then, are clearly instances of topic introduction or promotion, following Lambrecht's (1994: 176-177) definition, with the observation that an introduced element does not always have to be picked up in the following sentence. Another interesting point that becomes clear from Lambrecht's definition is that the introduced element does not have to be information-structurally new in order to be

promoted to topic status (*ibid.* 183). Once this pattern with late subjects becomes more prominent because of the loss of other types of inversion, an initial function of introducing an information-structurally new referent can easily expand to topic promotion a non-new referent. Some examples that suggest that such a function is developing are (54) and (55).

(54) *At her reward dined my cousen Fraunces and my cousen Magret.*
(edward-e1-h,363.285)

(55) *In the quier is an high tumbe of one of them porturid with his wife.*
(Ieland-e1-h,72.66)

The subject of (54), *my cousen Fraunces and my cousen Magret*, is ANCHORED, and the subject of (55), *an high tumbe of one of them porturid with his wife*, is OLD. These subjects, then, are not new but are nevertheless presented as the focus of the sentences: in (54), this is clear because the anchored element itself is new to the text, in (55), it seems to be more important that a new relation is presented between the topic and comment rather than a presentation of entirely new elements.

This analysis of the information-structural properties of preverbal and postverbal subjects in E1 shows that a general principle of given-first and new-last still plays a role in subject placement in E1, but that in many contexts syntactic factors have begun to dominate over information-structural factors. Most importantly, a syntactic default position for subjects before the finite verb is emerging, while the presence of an unaccusative verb sometimes seems to override information-structural factors. Where information-structural factors do play an important role is in the late subject construction, with a specific use of this position for topic introduction or promotion becoming dominant. In sum, the syntax limits the options for information-structural freedom for rearranging elements and apart from the late subject construction, the relevance of information status as a factor in subject placement is decreasing during the EModE period.

5.2.2.4 Conclusion

This section has investigated the loss of inversion after clause-initial PPs in the EModE period, aiming to establish the speed of this loss within the overall development of the loss of V2 and to determine whether the interaction between syntactic and information-structural factors in subject placement has changed since OE. I have shown that surface inversion is still at quite a high rate in E1, but that there is a steep decline in the later EModE periods. This means that the subject position in Spec,TP before the finite verb in T is becoming the default subject position, and the additional subject position in Spec,FP and the position for the finite verb in F are lost. At the same time, the PP-initial clauses continue to occur with

inverted subjects, even towards the end of the EModE period. I have identified several factors that influence this type of inversion, representing the emergence of the PDE locative inversion pattern: unaccusative verbs and presentational or topic introduction patterns favour inversion. Overall, it seems that the balance between syntactic and information-structural factors has shifted, with a syntactic default taking the lead and information-structural functions making use of specific syntactic options that are exceptions to the syntactic default.

5.2.3 Remnant inversion patterns in Modern British English

With inversion for all verbs already down to less than ten percent for all verbs except unaccusatives by the end of the Early ME period, this section turns to the MBE period, aiming to establish whether there remain any exceptions to the PDE patterns of inversion.

The selection of clauses is similar to the selection in the previous study, with one important difference: because no coding for unaccusatives is available in the PPCMBE, the queries cannot initially distinguish between types of verbs; this factor will be considered separately. The other selection criteria are similar: only PP-initial main clauses (IP-MAT) with a finite verb and a non-pronominal subject (so excluding pronominal and expletive subjects) were selected, and both conjunct and non-conjunct clauses were included. In addition, two types of clauses showing inversion were excluded, illustrated in (56) and (57).

(56) *And in that day did the Lord God of Hosts call to weeping, and to mourning, and to baldness, and to girding with sackcloth:*
(b1, burton-1762,1,1.4)⁷⁴

(57) *In no case is a monitor suffered to teach or tell the boys in his circle what the error is, unless they should all be equally ignorant:*
(b2, lancaster-1806,60.378)

Example (56) shows a clause with the finite verb *do* preceding the subject. The 6 examples showing this type of inversion were excluded, on the grounds that the introduction of *do*-support is a separate phenomenon (see Kroch 1989, Fischer et al. 2000: 134-135). Example (57) and 4 similar examples were excluded because they show subject-auxiliary inversion, meaning I-to-C movement, which is a different process from both the old V2-type of inversion and locative inversion.

Table 5.7 presents the results for remaining examples showing surface inversion in the PPCMBE, i.e. only distinguishing between subjects occurring before and after the finite verb.

⁷⁴ In the PPCMBE examples, the year of the text is given in the file name and I have added the subperiod to which they belong, with B1 representing the years 1700-1769, B2 1770-1839, and B3 1840-1914.

Table 5.7 Inversion of nominal subjects in PP-initial main clauses in the PPCMBE (conjunct and non-conjunct clauses combined)

	Inversion	Total	% Inversion
B1	17	318	5.3
B2	31	475	6.5
B3	31	469	6.4
Total	78	1262	6.2

$\chi^2(2) = .61.68, p = .74$

Table 5.7 shows that throughout the MBE period, inversion is only a marginal pattern: the percentage of surface inversion has dropped well below 10 percent for all subperiods. Interestingly, the percentage for B1 (5.3%) already represents a considerable decline from the percentage for E3 in Table 5.1, which was still at 13.8% for all types of verbs. This could possibly represent the differences between the two corpora, rather than a change specifically between the seventeenth and the eighteenth century, although the general decrease is in line with the development shown by the subperiods in each corpus.

The sentences in this selection that do still show inversion can quite straightforwardly be categorized according to the PDE observations about (locative) inversion. First, the overwhelming majority of these clauses have an unaccusative verb (86.5%, 75 out of 78 examples); this means that in the entire PPCMBE there are only three examples that show inversion in a non-unaccusative context. The unaccusative verbs come in three types. First, 26 examples have an unaccusative verb as the main verb, as illustrated in (58) and (59).

(58) *Over this hanas my smock, of a fine white silk gauze, edged with embroidery.*
(b1, montagu-1718,119.481)

(59) *To the right of these appears the Buet, a calcareous mountain.*
(b2, ruskin-1835,1,16.411)

The second category is represented by passive verbs (33.3%, 25 out of 75), as illustrated in (60).

(60) *With this was connected such a View, both of the Majesty and Goodness of God, as caused him to loathe and abhor himself, and to repent as in Dust and Ashes.*
(b1, doddridge-1747,33.271)

Note that the subject in (60) is unambiguously late, with the subject following both the finite verb and the participle; in total 15 of the 25 passive examples are unambiguously late. Another type of unaccusative verbs is represented by active

clauses where *be* is the main verb, as in (61), accounting for 24 of the 75 unaccusative verbs (32.0%).

- (61) *In this Posture were Affairs at the Inn when a Gentleman arrived there Post.*
(b1, fielding-1749,3,5.335)

These categories of unaccusative verbs still occur with inversion in PDE and so these MBE examples already show the situation that still exists in PDE. In other words, with regard to inversion, it seems that few changes have taken place since the MBE period.

The second factor for which the inversion examples are similar to locative inversion in PDE is the type of PP: locative PPs and local anchoring PPs dominate the types of PPs in the database. Examples (62) and (63) illustrate the category of locative PPs.

- (62) *On each side of this central mass are deposited huge folia of granite, those nearest to the centre being vertical, or nearly so, the next more sloping, the slope increasing as the folia retire from the centre.*
(b2, ruskin-1835,1,15.393)
- (63) *Under all roofs of this distracted City is the nodus of a drama, not untragic, crowding towards solution.*
(b2, carlyle-1837,1,159.590)

The other type of PP that stands out is the type illustrated in examples (64) and (65) below, where the PP is clearly linked to the previous discourse.

- (64) *Husbandry consists of many Parts [...]: All of them require Knowledge and Care in the Management, but some of them more Labour and Expeince than others, and some less;
of the last Sort is the Husbandry of Bees:*
(b1, axwell-1747,9.5-8]
- (65) *It is a strange museum, a revolting raree show for every chance comer to peep into, and wonder at, and come away from, complaining that he came before breakfast and has lost his appetite. The old bones might be let lie uncrunched by the thoughtless feet of staring hundreds; the corpses might be let sleep unoffended by their gaze.
Above this melancholy habitation rise the crags of Mont Mort, loaded with masses of thick snow.*
(b2, ruskin-1835,1,25.620-3)

In (64), the PP selects a subset of what has been mentioned in the previous sentence, and in (65), the PP refers back with a general description to the scene painted in the preceding sentences. The two types of PPs, locative and anchoring, relate directly to the functions of inversion identified for PDE, both the presentational function identified by Bolinger (1977) as well as the information-structural function specified by Birner & Ward (1998). Example (65) clearly ‘presents’ *the crags of Mont Mort*, bringing them onto the scene, and in a similar way example (62) presents *huge folia of granite*. In addition, all PPs in examples (62)-(65) are discourse-old, meaning that they all obey Birner and Ward’s relative constraint on the order of information, with given information preceding new(er) information.

A final observation that is relevant here is that there is very little evidence for the old type of inversion and the examples that do show this order seem to be deliberate archaisms. Examples (66) and (67) illustrate.

(66) *and **of these** was the whole earth overspread.*
(b3, erv-old-1885,9,1G.357)

(67) ***After these Things** came Jesus, and his Disciples into the Country of Judea;*
(b1, purver-new-1764,3,20J.173)

In both examples the verb is likely to be in F, with the subject following it in Spec,TP, based on the fact that in example (66), the subject precedes the passive participle and in example (67), it precedes the PP *into the Country of Judea*, which is presumably also VP-internal material. 8 of the 14 examples that show this order occur in Bible translations, and are for that reason likely to be archaic. Other examples also seem to be deliberate archaisms:

(68) ***In this manner** are all things well conducted, as the simplicity residing in the divine understanding produces that invariable order of causes;*
(b2, boethri-1785,171.399)

(69) *for **in Thy sight** shall no man living be justified.*
(b3, pusey-186X,294.210)

Example (68) is another translation of a formal text, Boethius’s *Consolation of Philosophy*, while example (69) is from a set of sermons (from Edward Bouverie Pusey), which are likely to rely heavily on biblical language. These examples which show the old V2-type of inversion, then, seem to represent archaisms rather than productive syntactic patterns during the MBE period.

In conclusion, by the MBE period, inversion has become a minority pattern and the sentences that do still show inversion can be categorized with reference to the PDE pattern of locative inversion while others seem to represent archaisms already in their time. Most importantly, the two PDE factors of unaccusative verbs and

locative or local anchoring PPs clearly stand out in the examples, while the function of inversion also seems related its PDE presentational function. While there is still some evidence for the old type of inversion, with the finite verb moving to a high position in the clause (F in OE terms) and the subject immediately following it, with other VP-internal material remaining within that VP, these examples are likely to represent deliberate archaisms at this period, occurring in bible translations and religious texts. Apart from this specific niche, inversion has by this time become limited to the locative inversion pattern that still exists in PDE.

5.2.4 Summary and conclusions

This section has investigated the loss of inversion of subject and finite verb in PP-initial clauses in EModE and MBE. The aim was to trace the decline in inversion in these sentences with respect to the loss of V2, while at the same time tracing the emergence of the locative inversion pattern that still exists in PDE. Another aim was to establish whether the interaction between syntactic and information-structural factors in subject placement have changed since the OE period. I have shown that the language developed a syntactic default position for the subject during the EModE period, while the late subject pattern—as locative inversion—now remained the only possibility for language users to manipulate the position of the subject for information-structural purposes.

The data for EModE showed that inversion was still quite frequent in the first subperiod, but that it soon became less frequent over the course of the EModE period, leading to a syntactic default position for the subject before the finite verb (in Spec,TP). With the development of this syntactic default, information-structural factors seem to have become less of a determining factor in subject placement, although the same general tendency of given-before-new as we saw in OE is still relevant. What has become more relevant for information-structural purposes than the variation between a high and lower subject position in the Left Periphery is the variation between the subject in Spec,TP and late subjects. In the EModE period, we see that the locative inversion pattern—with late subjects, locative PPs and unaccusative verbs—emerges as an important means of manipulating the order of arguments, as the other types of inversion are lost. This becomes even more prominent in the MBE period, with unaccusatives now accounting for nearly ninety percent of all inversion examples. By this time the old V2-type of inversion has become extremely rare, with the few examples that remain clearly representing deliberate archaisms. The MBE period, then, shows little difference from the PDE pattern of locative inversion. Overall, the data in these studies have shown that the syntactic changes lead to a limit on the information-structural freedom, while information-structural principles are expressed through the means that do still remain, such as late subjects.

Section 5.3 The loss of local anchoring

Changing the focus from the information-structural properties and position of the subject to the properties of the elements in clause-initial position, this section investigates the changes in the use of this clause-initial position from the EModE period onwards. It aims to establish the effects of the loss of V2 on the information-structural properties of the clause-initial elements that were discussed in Chapter 4, most importantly local anchoring by non-subject clause-initial elements. I show how subjects become more important as clause-initial elements towards the end of the EModE period and how local anchoring becomes a less important function of clause-initial PPs and objects throughout the EModE and MBE periods.

As became clear in Chapter 4, the use of the first position in OE is crucially different from the use of this position in PDE, something which was most clearly visible in the function of local anchoring: providing a link to the immediately preceding discourse by an information-structurally unmarked (i.e. non-contrastive) element in clause-initial position. At the same time, the clause-initial elements could also express marked or contrastive elements, although those categories were clearly less frequent than local anchoring. At the same time, we know that PDE has a preference for subjects in clause-initial position and that elements that do occur before the subject are often marked or contrastive, or have a specific text-structuring function (see Virtanen 1992). The question is when the character of the clause-initial position changes, especially with respect to the loss of V2. In the case of PPs, it is especially interesting to see how the time for the loss of local anchoring relates to the time of the loss of inversion in these clauses.

The changing use of the first position, or the loss of local anchoring more specifically, is relevant on two levels for the central question in this thesis: on the one hand it represents the loss of rearranging alternatives and restrictions on word order, which might in turn lead to a greater reliance on the passive for those functions. On the other hand it entails restrictions on the use of the clause-initial positions, in that it is more often the subject that occurs in clause-initial position, which in turn is likely to influence the functions of the subject (cf. Komen et al. 2014). The specific development in the clause-initial position as investigated in this section allows for a comparison between these information-structural consequences and the loss of inversion (syntactic V2), in turn putting us in a position to draw conclusions about the nature of V2 and the extent to which this specific use of the first position and the frequency of inversion are related.

This section is structured as follows. Section 5.3.1 aims to establish when English changes into a predominantly subject-initial language. Included in this study is an analysis of the other elements that occur in first position and an answer to the question which of these elements become less frequent as the subject becomes more frequent. Section 5.3.2 zooms in on clause-initial PPs and investigates changes in the information-structural character of clause-initial PPs compared to OE. Section

5.3.3 briefly comments on the function of objects in clause-initial position. Section 5.3.4 concludes.

5.3.1 The increasingly subject-initial nature of English

As a first step towards answering the question of what has changed in the use of the clause-initial position, this section investigates the frequency of different elements in this initial position. We saw in Chapter 4 that the subject is just one of the many options to start a sentence, besides objects, adverbs and PPs, while Los (2009) has proposed that in English the subject is the only unmarked theme, with presubject material having acquired a marked, often contrastive nature. This section aims to establish in detail during which period English developed this preference.

First, let us consider the frequency of subject-initial clauses as opposed to clauses beginning with a non-subject element. For this study I followed the same general criteria for the selection of clauses as in Chapter 4 and Section 5.2 of this chapter: all main clauses with a finite verb and a subject were selected, in both conjunct and non-conjunct clauses. Query 5.8 presents the details of the selection.

<i>Query 5.8 Subject-initial and non-subject-initial non-conjunct main clauses</i>
<p>a. IP-MAT idoms finite verb and NP-SBJ*</p> <p>b. IP-MAT idomsfirst NP-SBJ*</p> <p>c. IP-MAT idomsfirst CONJ* idoms NP-SBJ* <i>and</i> CONJ* iprecedes NP-SBJ*</p>
<p>Definitions:</p> <p>Finite verb:</p> <p>BEI BEP* BED* UTP *HVI *HVP* *HVD* *AXI *AXP* *AXD* *MD VBI *VBP* *VBD* *DOI *DOP* *DOD* NEG+BEI NEG+BEP* NEG+BED* NEG+AXI NEG+*AXP* NEG+*AXD* NEG+*MD NEG+VBI NEG+*VBP* NEG+*VBD</p> <p>Add_to_ignore: ** QTP META LATIN FW CODE LS</p> <p>QTP: quotations</p> <p>CODE: text mark-up information, page numbers</p> <p>META: metalanguage (e.g. stage directions)</p> <p>FW: foreign words</p> <p>LATIN: latin words</p> <p>LS: list markers</p>

Query line 5.8a selects all main clauses with a subject (labelled NP-SBJ*) and a finite verb, for which the definitions are given below the queries. Line 5.8b selects non-conjunct clauses, while line 5.8c selects conjunct clauses in which the conjunction is immediately followed by the subject. The final stipulation under 'Definitions' is that the technical categories QTP|META|LATIN|FW|CODE|LS were added to the add_to_ignore list, meaning that if one of these elements precedes the subject, the clause is still selected as subject-initial. To calculate the percentage of subject-initial

clauses, the results from lines 5.8b and 5.8c were combined and divided by the total number of clauses resulting from 5.8a.

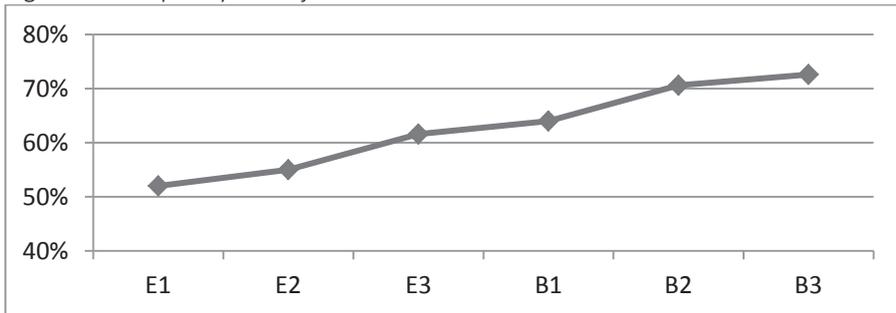
Table 5.8 and Figure 5.3 present the frequency of subjects in clause-initial position in conjunct and non-conjunct main clauses in Early and MBE.⁷⁵

Table 5.8 Frequency of subject-initial clauses in PPCEME and PPCMBE in conjunct and non-conjunct main clauses

	E1	E2	E3	B1	B2	B3
Subject-initial	8651	11893	11035	6087	9824	8549
Total no. clauses with a FinV and Sbj	16628	21627	17917	9504	13913	11780
% Subject-initial	52.0	55.0	61.6	64.0	70.6	72.6

EModE: $\chi^2(2) = 341.93$, $p = 0$; MBE: $\chi^2(2) = 193.85$, $p = 0$

Figure 5.3 Frequency of subject-initial clauses in PPCEME and PPCMBE



The table and figure show that there is an increase over time in the percentage of subject-initial main clauses with respect to all main clauses: from 52.0% in E1 to 72.6% in B3. The largest differences are between E2 and E3 (6.6%) and between B1 and B2 (6.4%), but more importantly, it is clear that it is a continuous development over time, with percentages steadily rising in all subperiods, rather than a switch at one particular point. If we consider the centuries represented by these labels for the subperiods, it means that the increase starts at the beginning of the sixteenth century, around 1500—the time usually taken to be the completion of the loss of V2, cf. Section 2.2.2—and continues until the beginning of the twentieth century. Although the last subperiod shows a smaller increase than the period before that, there are no indications that the development stops here.

⁷⁵ Conjunct and non-conjunct clauses behave in a very similar manner by this period.

The obvious follow-up question is what types of non-subject clause-initial elements become less frequent as the subjects become more frequent. I focused on the categories that are relevant as clause-initial elements in the OE V2 system: PPs, adverbs, objects. Two extra categories were added because they account for a large number of examples: the first category consists of clauses; the second category consists of nominal elements other than subjects and objects, such as vocatives and left-dislocated noun phrases. The remaining elements, such as verbal elements and adjectives, were grouped together as ‘Other’. An overview of all queries used is given in Query 5.9.

Query 5.9 Clause-initial adverbs, PPs, objects and clauses	
a.	IP-MAT idoms finite verb and subject
b.	Adverb IP-MAT idomsfirst ADVP* ELSE ALSO FP P+N IP-MAT idomsfirst PP* AND PP* idoms ADV ADV+P
c.	Clauses IP-MAT idomsfirst CP* IP* IP-MAT idomsfirst PP* AND PP* idoms CP*
d.	PP IP-MAT idomsfirst PP*
e.	Object IP-MAT idomsfirst NP-OB*
f.	Nominal IP-MAT idomsfirst NP* N
g.	Other IP-MAT idomsfirst QP* INTJ* RP ADJ* VAN BE* VB* NEG
add_to_ignore: QTP META LATIN FW CODE CONJ*	
Definitions: QP: quantifier phrase INTJ: interjection RP: adverbial particles FRAG: fragment NEG: negation	

Line 5.9b selects all adverbs, using the general label for adverbs, as well as the individual label for certain adverbs: ELSE for *else*, ALSO for *also*, FP for focus adverbs such as *only* and P+N for adverbs such as *indeed*. The second query under 5.9b selects a type of adverbs that we have already discussed before: adverbs that are coded as PPs, such as *therefore*. Line 5.9c selects clauses, coded in two different ways: either with a CP or IP as initial element, or as a PP immediately dominating a CP. Line 5.9d straightforwardly selects PPs, but needs to be executed on the complement file of the earlier queries to ensure that adverbs and clauses coded as PPs are excluded. The object and noun queries in 5.9e and 5.9f are straightforward. The rest category contains verbal categories (VAN | BE* | VB*), adjectives (ADJ*) and a number of other labels: QP* | INTJ* | RP | FRAG | NEG. The definitions of these elements are given below the queries. Note that I excluded the condition that the clause-initial element should immediately precede the finite verb or subject, because for this particular study doing so would result in a large category of clauses that would be added to the category of ‘Other’. Moreover, the focus is here only on

the clause-initial element, regardless of the order of subject and finite verb in those clauses which do not have a subject as the first element.

Table 5.9 shows the results of the queries above, giving the percentages of each type of clause-initial elements against the total number of clauses. This table, then, shows whether the different types of clause-initial elements remain stable throughout the periods, similar to the question for subjects in initial position. The remaining percentages represent the subject-initial clauses.

Table 5.9 Frequency of non-subject clause-initial elements with respect to total number of clauses (i.e. against the total number of clause-initial elements)

	E1	E2	E3	B1	B2	B3
ADVP	3270	3762	2052	933	859	909
% ADVP	19.7	17.4	11.5	9.8	6.2	7.7
PP	1554	2049	1438	960	1247	1078
% PPs	9.3	9.5	8.0	10.1	9.0	9.2
Clauses	1281	1503	1265	695	803	570
% Clauses	7.7	6.9	7.1	7.3	5.8	4.8
Object	206	255	198	73	80	62
% Object	1.2	1.2	1.1	0.8	0.6	0.5
Nominal	1124	1309	931	322	378	240
% Nominal	6.8	6.1	5.2	3.4	2.7	2.0
Other	540	856	997	432	721	371
% Other	3.3	4.0	5.6	4.6	5.2	3.2
Total no. of clauses	16628	21627	17917	9504	13913	11780

Table 5.9 shows that many categories become less frequent over the course of these periods at the time that the subjects become more frequent, with one category standing out: adverb-initial clauses show a steady decline, from 19.7% in E1 to 7.7% in B3, with an even lower percentage for B2, 6.2%. Nominal elements and objects also become continuously less frequent, with nominal elements starting at 6.8% and declining to 2.0%, and objects starting at 1.2% and decreasing to 0.5%. Note that objects are already a very small category in E1 and no longer represent an important category of clause-initial elements as they did in OE. Clauses also become less frequent overall, but their decrease in frequency only starts in the MBE period. The remaining categories, PPs and 'other',⁷⁶ are more or less steady throughout the

⁷⁶ A note is in order on the 'other' category, which, even if it accounts for only 3.2% to a maximum of 5.6% of the examples per period in Table 5.9, still represents a considerable number of examples, up to 997. The largest group in this category are sentences starting with an interjection. Adding this category to

periods. What this table shows, then, is that the subject becomes more frequent overall most clearly at the cost of adverbs and other NP elements, but more generally at the cost of almost all types of clause-initial elements, with the exception of PPs.

In addition to considering the frequency of each type against the total number of clauses, let us consider their frequency against the total number of non-subject clause-initial elements, i.e. excluding the subjects. Table 5.10 shows the type of clause-initial element as a proportion of only the non-subject clause-initial elements, i.e. the distribution of each element within the group of non-subject clause-initial elements.

Table 5.10 Frequency of non-subject clause-initial elements with respect to total number of non-subject-initial clauses

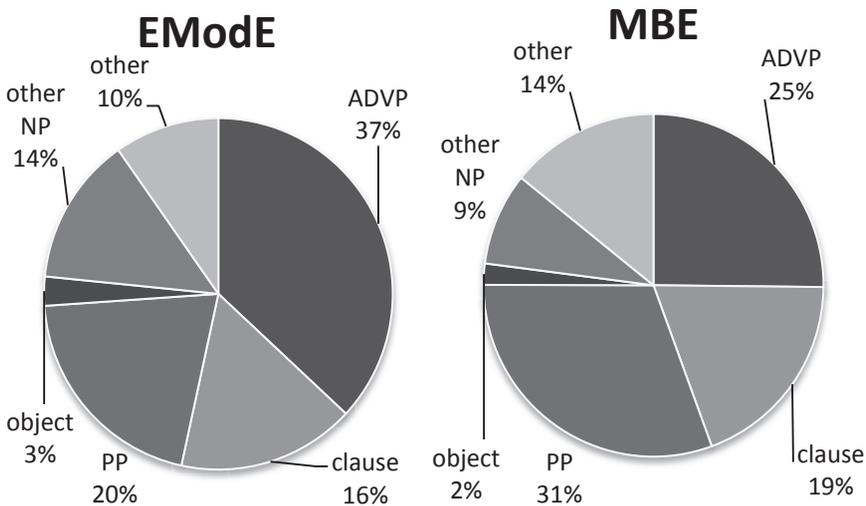
	E1	E2	E3	B1	B2	B3
ADVP	3270	3762	2052	933	859	909
% ADVP	41.0	38.6	29.8	27.3	21.0	28.1
PP	1554	2049	1438	960	1247	1078
% PP	19.5	21.0	20.9	28.1	30.5	33.4
Clauses	1281	1503	1265	695	803	570
% Clauses	16.1	15.4	18.4	20.4	19.6	17.6
Nominal	1124	1309	931	322	378	240
% Nominal	14.1	13.4	13.5	9.4	9.2	7.4
Object	206	255	198	73	80	62
% Object	2.6	2.6	2.9	2.1	2.0	1.9
Other	540	856	997	432	721	371
% Other	6.8	8.8	14.5	12.7	17.6	11.5
Total	7975	9734	6881	3415	4088	3230
% Total	100.0	100.0	100.0	100.0	100.0	100.0

The table shows that PPs come to make up a more considerable percentage of the non-subject clause-initial elements, from 19.5% in E1 to 33.4% in B3. Apart from that, most categories show a decline: adverbs drop from 41.0% in E1 to 28.1% in B3 (with an even lower percentage in B2, 21.0%); nominal elements go down from 14.1 to 7.4; and objects go down from 2.6% to 1.9%, also showing a steady decline over all periods. The only two categories that do not show a decline are 'Clauses' and 'Other', both of which do show fluctuating numbers over the course of the

the elements that can be ignored by the programme changes the percentages somewhat (especially for the 'other' category of course) but does not affect the trends.

subperiods but no diachronic development. In order to determine the overall development throughout the centuries, the following pie charts present the average of these percentages for each individual corpus, EModE and MBE.

Figure 5.4 Proportions of non-subject clause-initial elements in EModE and MBE



The pie charts highlight the decreased importance of adverbs in clause-initial position and the increased importance of PPs in clause-initial position: while adverbs are still at 37% in EModE, they are down to 25% in MBE; PPs are only at 20% in EModE, but are up to 31% in MBE. Two other categories now also show an increase: 'other' from 10% to 14% and clauses from 16% to 19%. In both periods, PPs, clauses and adverbs are the largest category, but the balance shifts from adverbs as the most frequent elements to PPs: they become the most prominent member among the non-subject clause-initial elements.

In conclusion, during the EModE and MBE periods the subject gradually gains ground as the clause-initial constituent at the expense of most other elements, except PPs, until around 70% of the clauses start with a subject. Adverbs in particular become less frequent in clause-initial position, but other NP elements as well, while objects are already infrequent in the first subperiod of the EModE period. At the same time, PPs in clause-initial position remain as frequent as they were, but the decrease in other elements in that position means that PPs gain importance as clause-initial elements.

5.3.2 Clause-initial PPs: a decrease in local anchoring and an increase in contrastive and frame-setting uses

Having established the changes in frequency for each of the clause-initial elements, this section addresses the question whether these changes are accompanied by functional changes. More specifically, it addresses the question whether the overall function of local anchoring—providing unmarked links to the immediately preceding discourse—and the smaller category of contrastive and frame-setting PPs are affected by these developments.

5.3.2.1 Loss of *P+D/PRO* in clause-initial position

The studies in Chapter 4 (Section 4.4.2) revealed that the majority of the PPs functioned as local anchors, providing an anaphoric link to the preceding discourse. This became most clear from the fact that roughly 40 per cent of the clause-initial PPs in *Orosius* and *Lives of Saints* consisted of only a preposition and a pronoun or demonstrative element, so a necessarily discourse-old element. Given the findings in the previous section about the increased importance of clause-initial PPs, but also given what is known about clause-initial PPs in PDE, the question is when the PPs with typical local anchoring elements as clause-initial elements are lost.

For this study, I selected non-conjunct main clauses with a finite verb, a subject, and a PP in initial position. As in the previous section, I excluded from the selection clauses or adverbs which are coded as PPs (see earlier queries for details). All texts from the PPCEME (so the Helsinki part, as well as the later additions) and PPCMBE were used. The next step in the selection was to distinguish the PPs in terms of the formal clues for anaphoricity, for which the procedure was similar to the one described in Section 4.4.1. The first category consists of PPs containing only a preposition and a demonstrative or pronoun, which are in all likelihood discourse-old. The second category consists of PPs which contain a demonstrative or pronoun, as well as additional material. The presence of the pronoun or demonstrative means that there is a good chance that these PPs are accessible to the reader. The last category consists of PPs with no formal indications for anaphoricity. The following examples from the PPCEME illustrate the three categories.

- (70) *With that* the sparkes appered euen as they had done of yore,
(stevens-e1-h,14.148)
- (71) *At this question*, al y=e= lordes sat sore astonied, musyng much by whome
thys question should be ment, of which euery man wust himselfe clere.
(moreric-e1-h,47.93)
- (72) *At S. Andres Akeland* the Dene of Akeland hath a great house: especially for
barnes and other houses of husbandry.
(leland-e1-h,75.140)

In example (70), the PP *with that* consists only of a preposition and a demonstrative, while in the PP *at this question* in example (71), the preposition is complemented with both a demonstrative and a noun. The clause-initial PP in (72) does not have an anaphoric element, consisting only of a preposition and a proper name.

One category was added for PPs that contain a definite article, which OE did not have, with demonstratives performing both the function of definite article and demonstrative pronoun (see Breban 2012; Denison 2006; McColl Millar 2000). Definite articles are coded as demonstratives (D) in the PPCEME and PPCMBE, but their information-structural nature is clearly different from that of demonstratives: while demonstratives are typically discourse-old, definite articles are more versatile, as the following examples illustrate.

(73) *And in the myddys of the Tower ys the place wher our blyssyd Savyor Crist Jhu ascendid vnto hevyn. Videntibus illis, etc.*

In the same Tower ys the ston vpon the whiche ower Savyor standing ascendid in to hevyn , in the whiche stone The prynte Of hys holy foote yett appere, And specially of the ryght foote.

(torkingt-e1-h,30.108-110)

(74) *In the beginnyng was the worde,*

(tyndnew-e1-h,1,11.6)

(75) *In the paroch chirch of S. Brandon at Branspeth be dyvers tumbes of the Nevilles.*

In the quier is an high tumbe of one of them porturid with his wife.

(Ieland-e1-h,72.65-66)

(76) *In this south isle, as I hard, was buried the grauntfather and grandedam of Rafe Raby , and they made a cantuarie there .*

In the waul of this isle appere the tumbes and images of 3. ladys, whereof one hath a crounet, and a tumbe of a man child, and a flat tumbe, varii marmoris.

(Ieland-e1-h,76.169-71)

The NPs in which these definite articles occur all represent a different information status. In (73), it is clear that the NP is discourse-old: *the same Tower* refers to the tower mentioned in the previous sentence. In (74), the NP is accessible but not discourse-old: there can only one beginning of the world. In (75), *the quier* is inferable from the mention of the parish church in the previous sentence, since churches usually have choirs. In (76), finally, the NP *the waul of this isle* is anchored through the linking of *waul* to *this isle*, which in turn is mentioned in the previous sentence. These examples show the versatility of the definite article in relation to

anaphoricity, and at the same time show that none of the elements are completely new and unanchored.

Query 5.10 sums up the queries that were used to select the different categories of PPs.

<i>Query 5.10</i> Clause-initial PPs	
On clause-initial PPs from previous query:	a. PP* idomsonly D* PRO
On complement file of query a:	b. PP* idoms NP* <i>and</i> NP* idoms D* <i>and</i> D* idoms the
On complement file of query b:	c. PP* idoms NP* <i>and</i> NP* idoms D*/PRO
On output file of query c:	d. PP* idoms NP* <i>and</i> NP* idoms D* <i>and</i> D* idoms a an A An

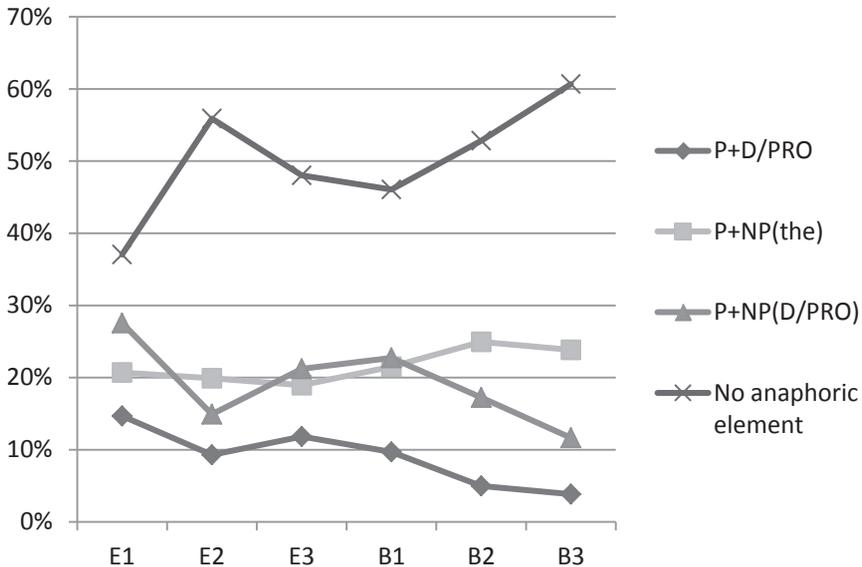
Line 5.10a selects PPs containing only a demonstrative (D) or pronoun (PRO), while 5.10b and c select PPs containing a definite article, a demonstrative or a pronoun. Finally, 5.10d selects NPs with demonstratives that are in fact indefinite articles; these are coded as D, like definite articles and demonstratives. The NPs with indefinite articles were added to the category of ‘No anaphoric element’. Note that following the ordering of queries presented here—query 5.11c is executed on the complement file of 5.10b—a PP with a demonstrative as well as a definite article will be included in the definite article category. Table 5.11 and Figure 5.5 show the results of these queries.⁷⁷

Table 5.11 Clause-initial PPs categorized according on the basis of anaphoric elements

	E1	E2	E3	B1	B2	B3
P+D/PRO	105	112	83	64	46	29
% P+D/PRO	14.7	9.3	11.8	9.7	5.0	3.8
P+NP(the)	148	240	133	142	230	180
% P+(NP)the	20.7	19.9	18.9	21.5	24.9	23.8
P+NP(D/PRO)	197	180	149	150	159	88
% P+NP(D/PRO)	27.6	14.9	21.2	22.7	17.2	11.7
No anaphoric element	265	674	337	304	487	458
% No anaphoric element	37.1	55.9	48.0	46.1	52.8	60.7
Total	715	1206	702	660	922	755
% Total	100.0	100.0	100.0	100.0	100.0	100.0

⁷⁷ If the conjunct clauses are included, the percentages become slightly different, but the overall trends remain similar.

Figure 5.5 Clause-initial PPs categorized according on the basis of anaphoric elements



The table and figure show two clear developments over time: a decline in use of PPs containing only a preposition and a demonstrative or pronoun (the bottom line); and an increase in the use of PPs that do not have any indication in terms of anaphoricity (the top line). Already in E1, at 14.7%, the percentage for the *P+D/PRO* category is quite a bit lower than the OE data, where it was 26.4%, and by the last subperiod of the MBE period, the percentage is down to 3.8%. The graph shows that the decline of this category is a more or less continuous development over time, with the exception of E3, where the percentage increases from 9.3% to 11.8%. What the graph also shows is that this decline of *P+D/PRO* examples seems to be to the benefit of the PPs with no anaphoric elements, because this is the only category that shows a more-or-less consistent increase over the subperiods. It is interesting that the development moves towards the category without anaphoric elements, rather than towards one of the middle categories, which would still represent linking, but of a weaker type. The two categories that do have one of these elements in addition to a noun and possible other elements show fluctuations over time, but there is no clear trend. That being said, *P+NP(the)* category shows an increase from E1 in comparison to B3, while *P+NP(D/PRO)* shows a decrease over the entire corpora. Overall, it seems that the major development for the decrease in *P+D/PRO*, and consequently, the increase in elements with no anaphoric elements, takes place in the MBE period.

There is one category of PPs which illustrates the development from standard local anchoring phrases to PPs with no indication for anaphoricity particularly well; this is a category of anaphoric adverbs that occur as the only element in some PPs, as illustrated in example (77).

- (77) [IP-MAT [PP [P *from*] [ADVP [ADV *thence*]]] *we went to the howse wher the Synnys of Mary Mawdley n war for govyn*.
(torkingt-e1-h,31.127)

These PPs were easy to select with a simple query⁷⁸ and the following table shows their decline over the EModE and Modern English period.

Table 5.12 Frequency of P+thence PPs in the PPCEME and PPCMBE

	E1	E2	E3	B1	B2	B3
P+thence	18	12	17	5	0	0
Total PPS	715	1206	702	660	922	755
% P+thence	2.5	1.0	2.4	0.8	0.0	0.0

While there are still a considerable number of examples in the EModE period, this type of PP has all but disappeared from the language in the Modern English period.

In conclusion, what these data show is a decreased presence of anaphoric elements in the clause-initial PPs, which suggests that local anchoring becomes a less important, or less prominent, function of the initial position. What is especially interesting with respect to the relation to the syntactic aspects of the loss of V2 is that the major changes only take place towards the end of the EModE period and in the MBE period. In terms of the actual dates, this means that it is only after roughly two centuries after the general loss of V2 that the use of the clause-initial position changes, during the period that inversion is already a clear minority pattern.

5.3.2.2 A study of the information-structural character of the clause-initial PPs in E1

Admittedly, the lack of indications for anaphoricity in the complements of the prepositions in clause-initial PPs does not necessarily mean that these PPs are indeed discourse-new; in the OE data one-third of the PPs that did not have an anaphoric element were 'given' (Table 4.24, Section 4.4.2). With the category 'no anaphoric element' larger in the EModE and MBE periods, it becomes even more important to gain insight into the information status of these PPs. This section therefore investigates the information status and local anchoring properties of clause-initial PPs. As in section 5.2.2.3, I focused on the E1 subperiod, because this is

⁷⁸ PP* idoms ADVP *and* ADVP idoms henc*|hens*|thenc*|thens*|here|there

the first period where we might expect changes, while some of the old patterns will also still be visible.

For this study, I only used the Helsinki part of the PPCEME corpus. Most of the selection criteria were similar to previous studies in this chapter: only non-conjunct PP-initial main clauses with a subject and a finite verb were selected and, for the sake of consistency, the PPs should immediately precede either the subject or the finite verb. The adverbs and clauses that are coded as PPs were excluded. The selection is similar to Section 5.2.2.3, with the difference that no selection was made with respect to the type of subject, so nominal, pronominal as well as expletive subjects were included. I also did not distinguish between inversion and non-inversion, because the focus is solely on the clause-initial PPs.

The result of the query was a selection of 204 clauses, from which I excluded 13 for one of the following reasons. First, there are a number of examples that turned out to be clauses after all, as in (78). Second, some PPs do not have a clear preposition+complement structure, as illustrated in (79). And third, some clauses represented speech, as in (80). These were excluded because it would be too difficult to determine their information status, since for these examples it is not only the text which is relevant but also the extra-textual factors.

(78) [PP [P *as*] [PP [P *yf*] [CP-ADV *a man would ryde for cause of helth*]]], *he desyreth not so much the mouing to ryde, as the effect of his helth.*
(boethco-e1-h,77.379)

(79) *Off our cher and well entretying at the rodys, And what Comfort was Don to vs, and Speciall that was seke and desesyd, by Sir Thomas Newporte, And Mayster William Weston, And Syr John Bowthe, and aftyrward by other Jentylmen of Englund ther, it war to long to wrytte.*
(torkingt-e1-h,57.317)

(80) *For your good wyl, I doo thanke your mastership moste hartelye,*
(mowntayne-e1-h,204.133)

Local anchoring was defined in Chapter 4 as clause-initial elements which are discourse-old, whose antecedent is in the immediately preceding sentence, and which are not information-structurally unmarked. The first aspect of the local anchoring function of the clause-initial PPs is the information status of the PPs. I used the same information status categories as in the previous studies (see Table 5.5). There were two problematic types of PPs in the current database. First, the category that already stood out with respect to inversion, that of opening lines of letters, presents problems with respect to the analysis of information status.

(81) *To my right worshipfull and my especiall good father Sir Robart Plompton, kt. be thes delivered.* (wplumpt-1510-e1-h,220.3)

The phrases are at the same time accessible and new: they are accessible because the writer and the intended reader know the person, but at the same time anyone who needs to deliver the letter may not know the person, and, more importantly, at the beginning of a stretch of discourse, any information is NEW or at best ANCHORED. The 10 examples that had this structure were excluded from the results. The second problematic category for the information status analysis is illustrated in (82) and (83).

(82) *At laste, I toke the waye to the towne,*
(mowntayne-e1-h,208.223)

(83) *Yn the meane tyme, I toke my waggon*
(mowntayne-e1-h,215.429)

In these examples, again, there is an odd combination of accessible and new information: while the concept will in all likelihood be inferable to the reader, it is completely new in the discourse and in addition there is no clear referent in these PPs, even though they contain an NP. The 8 clauses whose clause-initial PP consisted of such a phrase were excluded from the database. The exclusion of these two types of PPs results in 171 examples.

Table 5.14 shows the frequency of the different information status categories of the remaining PPs in E1.

Table 5.14 Information status of clause-initial PPs in E1

	<i>n</i>	%
OLD	111	64.9
ACCESSIBLE	22	12.9
ANCHORED	27	15.8
NEW	11	6.4
Total	171	100.0

The table shows that the majority of clause-initial PPs are OLD (64.9%), while only a small percentage are NEW (6.4%). The ACCESSIBLE and ANCHORED categories are roughly equally frequent, 12.9% and 15.8%, respectively. If we only consider the binary distinction of ‘given’ (OLD+ACCESSIBLE) versus ‘new’ (ANCHORED+NEW), the dominance of given elements in clause-initial position becomes even clearer, with roughly three-quarters ‘given’ (77.8%), and just under a quarter of the examples ‘new’ (22.2%). However, despite the dominance of ‘given’ elements, the category of ‘new’ elements is quite considerably larger than it was in the OE study, where only 8.3% of the clause-initial PPs were analysed as ANCHORED or NEW. What these data show, then, is that the information status of the clause-initial PPs has changed, while at the same time, there is still a clear tendency for clause-initial PPs to be given. This tendency is

not surprising in light of the strength of the general principle of given-before-new order of information. Even though the balance between marked and unmarked themes in initial position (cf. Lambrecht 1994: 31-32) may change, it is unlikely that a language will abandon this principle completely.

The second aspect of local anchoring is the distance to the antecedent if the referent is discourse-old. In the OE data, only 7.5% of the clause-initial elements were found to not have such a link to the previous discourse, and of the anaphoricity scores that were measured, the antecedent was in most cases in the previous sentence. For this analysis, I used the same categories as in Chapter 4, using a window of five main clauses automatically selected by *CorpusStudio*. I added one category to the ‘general’ links, illustrated in the following examples.

- (84) *In this yere/ and .xx. daye of the Moneth of Nouembre was a great Counsayll holden at the Whyte Freres of London*
(fabyan-e1-h,174V.C1.209)
- (85) *In this chapter is declared the fiue thinges conteyned within the head.*
(vicary-e1-h,28.4)

The reference is to a preceding title, although this has not always been included in the corpus, and this type of reference is different from the other ‘general’ categories in that the reference is not strictly speaking to the content of the discourse, but only to the formal layout of the text. There are 8 examples like this in the database and I have analysed them as OLD, with a distance to the antecedent of 1 clause. Finally, any item that was contrastive was analysed as ANCHORED, as illustrated in the following examples.

- (86) *At another time, Vaughan saith, Throckmorton shewed him, that he had sent a Poste to Sir Peter Caroe to come forwarde with as muche speede as might be, and to bring his force with him.*
(throckm-e1-h,1,67.C1.220)
- (87) *vpon the Thursdaye folowyng* was put in to the comon House a byll deuysed by syr Iohn~ Bagot than Prysoner in the Towre]
(fabyan-e1-h,168V.C1.26]

The PPs *At another time* and *vpon the Thursdaye folowyng* are both contrastive to the event described in the previous sentences and are anchored to the previous discourse because of this contrast. Table 5.15 below shows the results of the analysis of the anaphoricity of the initial PPs. It differentiates between three categories: a link to the previous clause, a link to 2 to 5 clauses before the relevant clause, and no link.

Table 5.15 Linking of clause-initial PPs in E1

Distance to antecedent	Type of antecedent	n	%
1 clause	Object/subject	12	7.0
	Topic/title/speech/event	78	45.6
	Anchor (and contrast)	31	18.1
	<i>Subtotal</i>	121	70.8
2-5 clauses	Object/subject	13	7.6
	Anchor (and contrast)	5	2.9
	<i>Subtotal</i>	18	10.5
No antecedent	No link	32	18.7
Total		171	100.0

The table shows that PPs with an antecedent in the immediately preceding clause—either a specific antecedent or a ‘general’ reference—make up just over half of the clauses in the database: 52.6%. An additional 18.1% provide a link to the immediately preceding clause, but only by inference or anchoring. Nearly twenty percent (18.7%) of the clause-initial PPs have no link to the preceding discourse. The table clearly shows that if there is a link to the preceding discourse, this link is local: only in 10.5% of the examples is the antecedent in a clause before the immediately preceding clause. This trend of a high percentage of local links is comparable to the findings for OE, but as with information status, the category of examples that do not have a link to the preceding discourse is already slightly larger in this first EmodE subperiod. Overall, however, local links remain the dominant pattern, and links to more than one clause preceding are infrequent.

The last aspect of the local anchoring function is the information-structural character of the clause-initial PPs, i.e. whether they have a contrastive or frame-setting function. I analysed the examples in the database following the same definitions as in Chapter 4, Section 4.4.2: contrastive examples are those for which a contrastive element is present in the text, as in (88); frame-setting PPs, in addition to presenting the context for which the statement that follows will hold, show an implied contrast but the alternative option is not present in the text, as illustrated in (89).

(88) *At her reward [at a favoured place at table] dined my cousen Fraunces and my cousen Magret.*

At mine sate the French embassadour.

(edward-e1-h,363.285-6)

(89) *In that time saw I Maister Doctour Lattemer come in to the gardein,*
(morelet2-e1-h,503.15)

Any PP that does not have a code for contrast or frame-setting is taken to be information-structurally neutral, and, crucially, when it is discourse-old, is taken as an example of local anchoring, as in (90). A fourth category was used for examples that have no code for contrast or frame-setting, but are also not discourse-old, as in (91).

(90) *After these things* cam Iesus and his disciples into the Iewes londe,
(tyndnew-e1-h,3,20J.192)

(91) *From Farley* I ridde a mile of by woddy ground to a graung great and welle
buildid, that longid to Henton-priorie of Chartusians.
(Ieland-e1-h,139.233)

Table 5.16 presents the frequencies of the four categories.

Table 5.16 Information-structural function of clause-initial PPs in E1

	<i>n</i>	%
Contrastive	22	12.9
Frame	48	28.1
Local anchoring	91	53.2
Other	10	5.8
	171	100.0

The table shows that contrastive and frame-setting uses make up quite a considerable part of all PPs, with 41% of all PPs belonging to this category. While this specific aspect was not separately investigated in Chapter 4, it was clear from the other evidence that local anchoring was the main function of the initial PPs. This function of local anchoring is still important, accounting for just over half of the PPs, but in this first subperiod of EModE its importance is already decreasing.

A closer analysis of the examples shows that some of the patterns, especially those that can be characterized in terms of local anchoring, are limited to certain texts. For instance, one of the texts, a travelogue by Richard Torkington, contains all the *from thence* examples in the E1 period (see example (77) above). There are other types of clause-initial PPs which seem to be limited to one text. One of these texts is a chronicle by Robert Fabyan, which has many examples of time adverbials as clause-initial PPs, as in the following examples.

(92) *In this yere* the Seneshall of Henaude came into this Lande with a goodly
Companye of Henauders & other straungers
(fabyan-e1-h,173R.C1.172)

(93) *Vpon the .iiii. daye* came into the felde an Esquyre Henauder agayn whom
ranne the sone of sir Iohn Cheyny (fabyan-e1-h,173V.C1.179)

These examples in fact represent the only type of clause-initial PPs that are found in the Fabyan text (cf. earlier comments and Los (2013) on the use of V2 as episode boundary marker).

In sum, these data have shown that the clause-initial PPs in E1 to a large extent still show the same characteristics as the PPs in OE, but that the local anchoring properties seem to have become less prominent, and that the information-structurally marked categories, like contrast and frame-setting, have gained importance. The clause-initial PPs remain mostly discourse-old and the linking that takes place in initial position continues to be local: links are in the majority of cases to the immediately preceding sentence. The majority of clause-initial PPs are still information-structurally unmarked. However, what is most important in these data is that already at this early stage after the loss of V2—these data are, after all, based on early sixteenth-century texts—the character of the initial position is already showing signs of change.

5.3.2.3 Conclusion

The studies presented in this section have provided two very clear indications that the information-structural character of clause-initial PPs has changed: PPs which can be analysed as OLD or ACCESSIBLE on the basis of form decrease, while the percentage of PPs that can be analysed as having a local anchoring function also becomes lower, to the benefit of contrastive or frame-setting PPs. With respect to the timing of these developments, the data on the anaphoric character of clause-initial PPs show that the change takes place in the transition from the EModE to the MBE period, but at the same time, the analysis of the information-structural properties of clause-initial PPs in the E1 subperiod showed that signs of changes in use are already present in the early sixteenth century.

5.3.3 The increasing restrictions on object fronting

Having considered the PPs, let us look briefly at the other important category of clause-initial elements that was discussed in Chapter 4, fronted objects. The corpus study in section 4.3 showed that object fronting in OE represents an alternative to the information-rearranging function of the long passive: it allows an underlying object to occur in clause-initial position, like the passive, and is regulated by the same relative information-structural constraint as the long passive. In addition, the main function of initial objects was local anchoring. While objects were a relatively frequent element in clause-initial position in OE, the data in Section 5.3.1 already showed that objects have become very infrequent in initial position in the corpora: with respect to all main clauses, objects represent only 1.2% of all clauses in E1 and this percentage decreases to 0.5% in B3. This means that the development of objects in initial position is different from that of the clause-initial PPs: the questions of loss of local anchoring and loss of inversion are less relevant because the initial objects come to represent a marginal option in the language.

The decline in object fronting and the loss of inversion after objects has been investigated in quite some detail in the literature, for instance in Van Kemenade & Westergaard (2012), but most comprehensively by Speyer (2010). As discussed in Chapter 4, Speyer distinguishes between three types of object fronting (object topicalization) in PDE, which are all characterized by focus and/or contrast. For OE, he adds a type of object fronting which has a comparable function to local anchoring as defined in Chapter 4. Speyer then traces the decline in object fronting and also attempts to trace the decline of the specific local anchoring type of object fronting. For the general decline of object fronting, Speyer presents data from the historical corpora of English, up to the EModE period (his study was carried out before the PPCMBE was published), showing the number of topicalized direct objects against the total number of clauses with a direct object.

Table 5.17 Rate of direct object topicalization (from Speyer 2010: 27, his Table 1)

	OE 1-2	OE 3-4	ME 1	ME 2	ME 3	ME 4	EmE 1	EmE 2	EmE 3
sent. with DO	6184	10002	5329	3642	9608	5583	7719	10103	7057
Whereof topicalized	736	1080	570	228	558	257	376	428	247
% topicalized	11.9	10.8	10.7	6.3	5.8	4.6	4.9	4.2	3.6

Speyer's data show a continuous decline in direct object fronting from OE to EModE, from 11.9% to 3.6%, with the sharpest decline taking place from the M1 to the M2 subperiod (10.7% to 6.3%). This shows that object fronting in general declines, confirming the findings in Section 5.3.1, even though the numbers and percentages are somewhat different. Speyer compares the number of examples against the number of transitive clauses instead of all main clauses, which partly explains the difference in percentages between these numbers and my Table 5.9. In addition, the actual number of examples in EModE is higher in Speyer's study than in Table 5.9, which may be due to a difference in selection, for instance including subclauses and objects which are not clause-initial but still precede the subject. Speyer does not give his queries so a detailed comparison is not possible.

Speyer also presents data that can give us an insight into the loss of the local anchoring function of clause-initial objects. His main characterization of the additional type of object fronting specified for OE is that the fronted objects are aboutness topics, representing "entit[ies]" and that they are anaphoric elements, often pronouns, and non-contrastive (2010: 38), as illustrated in (94).

- (94) *þone asende se Sunu,*
 this sent the Son
 'The son sent this one'
 (coaelhom,ÆHom_9:114.1350; Speyer 2010: 38, his 17a)

Speyer's definition is slightly different from the notion of local anchoring as used in this thesis, because it is based on the notion of aboutness topic, but it nevertheless overlaps with this notion in some crucial aspects, most importantly the discourse-link provided by the object and the information-structurally unmarked character of this object. As an approximation of the decline in this type of object fronting, Speyer gives the following table, showing fronting of personal pronoun objects, this time against the total number of clauses with a pronominal object.

Table 5.18 Topicalization of personal pronoun objects (Speyer 2010: 39, his Table 3)

	OE 1-2	OE 3-4	ME 1	ME 2	ME 3	ME 4	EmE 1	EmE 2	EmE 3
All pr.obj.	200	603	285	213	454	316	107	155	96
Whereof topicalized	22	40	11	2	8	0	1	0	1
% topicalized	11.0	6.6	3.9	0.9	1.8	0	0.9	0	1.0

This table shows a loss of topicalization of personal pronoun objects, with very low percentages from ME2 onwards. Speyer claims that the late examples with pronouns are contrastive or archaic (2010: 40). An example of such a deliberate archaism is given in (95).

(95) *'Tis said, here lives a woman, close familiar to the enemy of mankind. Her I'll consult, and know the worst.*

(Charles Jennens: Libretto to Handel's oratorio 'Saul', 3rd act, 1739; Speyer 2010: 40, his 18)

Speyer's data locate the loss of this type of topicalization in the ME period, meaning that by ME, the types of topicalization that are found are the types that persist up to PDE, crucially without the local anchoring type. In short, while the local anchoring type is lost almost completely, topicalization in general shows a decline—already from the OE period onwards—but arrives at a stable percentage in the late ME period and remains to be used in later periods. Interestingly, Speyer proposes that the loss of topicalization is a result of the loss of inversion, because inversion in the relevant cases provided an option for reordering which obeyed a prosodic requirement, which he calls the Clash Avoidance Principle (2010: 61) (although cf. Van Kemenade & Westergaard 2012, who point out some problems with Speyer's scenario).

Let us investigate object fronting in the PPCEME and PPCMBE, in order to add to Speyer's data in the context of local anchoring, and to extend the data to include MBE. Query 5.11 presents the selection criteria.

<i>Query 5.11</i> Clause-initial objects
a. IP-MAT idoms NP-SBJ* <i>and</i> IP-MAT idoms NP-OB*
b. NP-OB* precedes NP-SBJ*
On output of query b:
c. NP-OB* idomsonly PRO*
d. NP-OB* idomsonly D*

Line 5.11a selects all main clauses which have a subject and an object, while 5.11b selects all examples in which the object precedes the subject. Lines 5.11c and 5.11d provide a further selection on the type of object—containing only a pronoun or only a demonstrative, respectively. Note that line 5.11b does not stipulate that the object needs to be clause-initial, only that it should precede the subject. Table 5.19 presents the results of the first two queries, giving the number of fronted objects against all main clauses with an object NP, and the number of fronted pronominal objects against all main clauses with a pronominal object.

Table 5.19 Fronted NP and pronominal objects in PPCEME and PPCMBE main clauses

	E1	E2	E3	B1	B2	B3
% fronted NPs	4.4% (262/5945)	3.6% (297/8339)	3.3% (141/6848)	2.3% (80/3545)	1.8% (85/4855)	1.6% (64/4085)
% fronted pronouns	0.7% (9/1250)	0.2% (4/1944)	0.3% (6/1780)	0.4% (3/760)	0.2% (2/964)	0.3% (2/677)

Table 5.19 shows that the percentage of fronted NPs is already low in E1, which is only to be expected after the general decline of initial objects as shown in Table 5.9. The percentages for EModE—from 4.4% to 1.6%—are comparable to Speyer’s percentages (Table 5.17, from 4.9% to 3.6%), although Speyer’s number of examples is higher. The MBE data show that this percentage continues to fall, until only 1.6% of the objects is fronted in B3. The more dramatic loss of pronominal object fronting is also visible from Table 5.19, with the percentage already at 0.7% in E1 and declining to 0.3% in B3. Again, these percentages are comparable to Speyer’s percentages (Table 5.18, between 0 and 1.0% in EModE), although this time the number of examples in my database is higher (this may be due to the fact that I included all pronouns and Speyer specifically states “personal pronouns”). Already in E1, the examples, as in (96), seem highly contrastive or archaic—corresponding to Speyer’s claim. This is even stronger for the remaining examples in the MBE, as in (97)-(99).

- (96) *For many he had,
but **her** he loued, whose fauour to say the trowth for sinne it wer to belie y=e=
deuil she neuer abused to any mans hurt, but to many a mans comfort &
relief:*
(moreric-e1-h,56.132-3)
- (97) *There was at that time with him one Sicinus, a Persian Prisoner, in great
Favour with Themistocles, insomuch that he taught his Children.
Him he sent privately to the King, to give him a true Information of the
intended Flight, advising him to send part of his Navy about the Island, which
encompassing the Grecians, would prevent their Escape,*
(b1, hind-1707,321.259-60)
- (98) *Coleman being sworn, and looking at them, and then pointing Batteah. That
Gentleman's Name is Batteah, and I have seen him at Derby. And t'other is
Chadwick,
and **him** I have seen at Manchester,*
(b1, townley-1746,46.568-71)
- (99) *Corsini has lost the only one he could have ventured to make pope,
and **him** he designed;*
(b1, walpole-174X,5,11.240-1)

In all these examples, the pronoun is not just anaphoric but seems to bear focus or emphasis, singling out one person among a range of options. In examples (96) and (98) this is especially clear: in (96), *her* is contrasted with the *many* in the previous line; in (98), there is a contrast between the two men that the witness points at, and who he has seen at different places.

Interestingly, the other type of objects selected by Query 5.11, independently used demonstratives, do not decline as dramatically and at the same speed as the rates given in Table 5.19. Table 5.20 presents the percentage of fronted demonstratives against the total number of clauses with an independently used demonstrative.

Table 5.20 Fronted demonstrative objects in main clauses in the PPCEME and PPCMBE

	E1	E2	E3	B1	B2	B3
% fronted	54.5%	55.9%	44.7%	43.9%	27.9%	31.8%
demonstratives	(55/101)	(85/152)	(63/141)	(25/57)	(17/61)	(14/44)

Table 5.20 shows that until the B1 period, almost half of these demonstratives are fronted. Only in the last subperiods do the percentages show a decrease, with lower percentages in B2 and B3 (27.9% and 31.8% respectively). Although these

demonstratives are likely to be discourse-old information (exceptions are possible of course, for instance in cataphoric references), like the fronted pronouns in the later periods, they are mostly used contrastively or represent archaic examples. Examples (100) and (101) present two examples from the E1 subperiod, which are clearly archaic: they show inversion of finite verb and subject, and occur in a bible translation.

- (100) *This sayde he, not that he cared for the pooer: but because he was a thefe, and kept the bagge, and bare that which was geven.*
(tyndnew-e1-p1,12,1J.16)
- (101) *This sayde Iesus signifyinge what deeth he shuld dye.*
(tyndnew-e1-p1,12,20J.68)

The following examples from MBE illustrate the contrastive use.

- (102) *but I thank God that he preserved our precious Alix, for **that** she is,*
(victoria-186X,1,151.378-9)
- (103) *I hope that my opinions and my actions will not be quoted in opposition to my Government, for **that** beloved Papa never permitted;*
(victoria-186X,1,207.476-7)

In (102), the demonstrative is emphasized to stress the quality of *Alix*. (103) presents an interesting case because the PPCMBE annotation contains a comment that the word *that* is “emphasized in [the] original” text. However, even in MBE, examples occur where the text contains little signs of contrastive use.

- (104) *As Francisco had predicted, the collector asked me whether I preferred gin or Champagne , and, on receiving my answer, brought in a pint bottle with a large label on its side- Fine Crab-apple Cider. **This** he opened,*
(reade-1863,224.588-90)

While there is certainly a contrast between two drinks in this passage, the actual demonstrative seems to simply provide an unmarked discourse-link between two sentences, and is not clearly contrastive; at this point in the text, there is not a contrast between two bottles. In terms of the syntactic structure, the sentences with a fronted demonstrative seem to represent a fixed pattern, all occurring in short sentences and followed by a pronominal subject and a verb. The fact that pronouns and demonstrative elements are almost invariably specific referents often leads to a contrastive or focused effect, even when they do not occur in initial position: it is always a matter of a choice among alternatives, a crucial aspect in Krifka’s (2007) definition of contrast. The initial position is a likely position for these short

contrastive elements to occur, even more so once the unmarked use is lost and the marked use dominates this position.

In conclusion, the number of objects in clause-initial position has decreased and the trends that Speyer already noted for EModE continue in MBE. What is decidedly different in the development of clause-initial objects from the development of clause-initial PPs is that the change in use is more drastic (no local anchoring at all) and that the timing is earlier: it is already in Middle English that there are indications that the local anchoring use of the clause-initial object declines, i.e. before V2 was well and truly lost.

5.3.4 Summary and conclusions

This section aimed to establish in what way the use of the clause-initial position changed after the ME period. It followed up on two aspects identified in the previous chapter for the clause-initial position: the function of local anchoring and the frequency of non-subject elements in clause-initial position. The data presented in this section show that local anchoring became a less important function of presubject material in the EModE and MBE periods and that subjects became more important as clause-initial element, locating the changes in the early MBE period.

Subjects become the most frequent element in first position, a development which takes place consistently from the E1 period onwards. The clause-initial elements that are most affected by this are adverbs and direct objects, while PPs remain equally frequent in initial position throughout these periods. At the same time, the importance of the PPs in first position increases because they become the most frequent member among the remaining non-subjects in initial position. In addition to these quantitative changes, there are qualitative changes in the function of the elements in first position: objects are overwhelmingly contrastive, while PPs can still be unmarked, but the category of frame setters and contrastive PPs has become larger. As a consequence, the subject became even more important in clause-initial position, because it still could be unmarked and discourse-linking, but now as the only element. The developments described in this and the preceding section put greater pressure on the subject, both in terms of information structuring (subjects are the only unmarked theme) as in terms of syntax (most sentences start with a subject).

Section 5.4 Conclusion

This chapter has investigated aspects of the loss of V2 in the EModE and MBE periods, after the date that has usually been taken as the loss of V2, 1500. The main aims were to establish the rates of the loss of inversion, and to trace the information-structural functions of V2 that were identified in the previous chapter. The most important conclusions from this chapter are that the presubject position lost its function of local anchoring while its contrastive and frame-setting uses became more prominent, towards the end of the EModE period and during the MBE

period. At the same time, the language developed a default position for the subject, before the finite verb, which left little room for manipulation of the order of elements in the clause for information-structural purposes, with the exception of the late subject position, which developed into the locative inversion pattern that is still in use in PDE.

While V2 is already a minority pattern at the end of the ME period (around 1500 AD), inversion was by no means lost in all contexts. For PP-initial clauses in particular, we see that the overall decrease from a majority pattern to a minority pattern only takes place during the EModE period. At the same time, inversion is from the first subperiod influenced by particular factors—PP and unaccusative verbs—which leads to a specific pattern emerging while overall inversion is lost. Although this is already prefigured in the earlier periods, with unaccusative verbs already allowing for late subjects in OE and having this particular presentational function, it is in this period that the locative inversion pattern as it is still used in PDE rises in prominence as the other patterns of inversion are lost.

In addition to the loss of inversion, changes take place in the function and information-structural character of the initial position, which are not directly related to the syntactic loss of V2. Most importantly, the function of local anchoring which was identified as the main function of the initial, presubject position in Chapter 4 loses its dominance, while the contrastive and frame-setting use of the presubject position, which was only a minority pattern in OE, increases in importance. Already in E1, this function makes up forty percent of the examples, a development which is likely to continue in the following centuries, as the data on the loss of anaphoric elements in clause-initial PPs already suggest. It is likely that this development starts in E1 but only gathers full force in the MBE period.

In conclusion, inversion after clause-initial PPs is lost later than what has generally taken to be the endpoint for the loss of V2, in the late EModE period, while locative inversion, with a strong information-structural motivation, emerges as a specific exception to the overall loss of inversion. During this same period, but most convincingly in the MBE period, main clauses come to be predominantly subject-initial. It is at this time that we also see the most convincing evidence of a change in the information-structural use of clause-initial PPs: the function of local anchoring decreases and the PPs become predominantly contrastive or are used as frame setters. The fact that these changes are only after the loss of inversion suggest that they are a consequence of rather than strictly speaking part of, the loss of V2.

6. Responses to the Loss of Verb Second

Section 6.1 Introduction

This chapter shifts the attention from the actual loss of verb second (V2) to the consequences of this loss and the syntactic and information-structural developments that accompany it, as identified in the previous chapter. Specifically, it investigates the relation between the development of the subject and the clause-initial position on the one hand and the development of passives, middles and non-agentive subjects on the other. Crucially, these constructions have all been described in the literature either as information-rearrangers or subject-creators. The studies in this chapter show that there is evidence for an increase in the use of all these constructions over the course of the Early Modern (EModE) and especially the Modern British (MBE) period. In addition, their functional relevance as subject creators and information structuring devices provides support for the hypothesis that this increase is a response to the loss of V2.

The previous chapter established dates for the loss of V2 in more detail for a number of specific contexts, most importantly with unaccusative verbs and PPs in first position. It also identified certain information-structural aspects of the loss of V2, namely the increasingly subject-initial nature of the language in EModE and MBE, as well as an increased contrastive, and hence *marked*, use of the clause-initial position. What still remains unclear is what the consequences are of such a development in the language and specifically, whether the corpora can provide sufficient evidence for previous claims that certain developments—such as an

increase in the use of passives—can be shown to be a response to the restrictions on the clause-initial position. The proposal that English developed certain constructions such as passives and middles to compensate for the loss of word order freedom has been made by for instance Kirkwood (1978), Hawkins (1986) and Halliday (2000), but this has been mostly based on general observations and there is little data available for the developments in the intervening periods.

Los (2009) and Seoane (2006) have connected several developments specifically to the loss of V2, instead of only to the restrictions on word order in Present-day English (PDE). They point to the ability of long passives to provide a response to the restrictions on word order options because of their rearranging function, as well as the ability of both long and short passives to move arguments to subject position—relevant because of the increased importance of subjects. The new passives that were introduced in the Middle English (ME) period represent a special class among passives. Another strategy identified by Los (2009) as a response to the loss of V2 is the use of non-agentive subjects (Rohdenburg 1974), and the use of middles (see Hundt 2007) can be seen as a similar strategy. While these developments have previously been identified as innovations by e.g. Hawkins (1986) and Kirkwood (1978), Los (2009, 2012) and Los & Dreschler (2012) have connected this to the loss of V2 and the need for subjects.

The aim of this chapter is to investigate the developments in passives, middles and non-agentive subjects after the ME period and establish whether these can be shown to be consequences of the loss of V2, or, more precisely, responses to the consequences of the loss of V2 as described in the previous chapter: an increasingly subject-initial language and a marked presubject position. Specifically, this chapter aims to establish whether (i) there is an increase in use of either of these constructions which takes place simultaneously to the developments described in the previous chapter; (ii) new strategies are introduced in this period; (iii) these constructions have or develop a function which relates to the need for more subjects or information-rearrangers.

This chapter is structured as follows. The first sections trace the development of the passive after the ME period: Section 6.2 investigates whether the use of the passive in general increases, as well as the long passive, while Section 6.3 provides a follow-up to the studies in Chapter 3 by examining the developments of the new passives after their introduction in the ME period. Section 6.4 moves the focus to middles and non-agentive subjects and investigates their structural properties as well as their diachronic development. Section 6.5 concludes.

Section 6.2 The development of the passive after ME

This section focuses on the development of the passive after the loss of V2 and aims to establish whether there is an increase in the use of passives in general and long passives in particular, both compared to the OE data in Chapter 4 and Seoane's (1999, 2000) data for EModE. The second aim of this section is to investigate the

information-structural function of passives. I present data which provide a further refinement of the previously proposed increase in the use of passives with respect to transitive actives, as well as an increase in use of long passives with respect to short passives. In addition, I show that the information-rearranging function of the passive becomes more important, while at the same time the language makes full use of the ability of passives to create subjects which in turn fit into the natural unmarked structure of sentences, beginning with a subject.

Two properties of the passive make it a potential candidate to act as a response to the loss of V2: its information-rearranging function and its subject-creating function. The information-rearranging function has been noted before by many authors, such as Halliday (1968) and Chafe (1970), cf. Section 2.3.2. In addition, Birner (1996) has provided corpus data to support the observation that the passive restores given-before-new order of information in the sentence in PDE, and Seoane (1999) has done the same for EModE. The data in Chapter 4 showed that this function was already present to some extent in the OE period, but to a more limited extent than in later periods. In addition, Los (2005, 2009) and Seoane (2006) have focused on the fact that the passive can create a non-canonical (i.e. non-agentive) subject by moving an argument to subject position. As pointed out by Los, this function becomes especially important in the history of English because of the increased need for subjects after the loss of V2, as described in the previous chapter.

If the passive indeed acts as a response to the loss of V2, there are two likely areas in which this might become visible: an increase in use of the passive and functional changes. In the literature such an increase has been described, for instance by Siewierska (1984), Halliday (2000) and Seoane (2000), but the evidence is not overwhelming. In particular, what is not available yet is a detailed overview of how the frequency of the passives changes after the EModE period. This is where the larger annotated corpora which are now available can play a role. It is not only the frequency of the passive, however, that provides evidence for a change in its role within the language, but there is evidence that its function changed too—the information-reordering function, present since the OE period, seems to become more pronounced. At the same time, the number of available constructions that can express the same function decreases: while V2 OE speakers had multiple strategies for information-rearranging, topicalization and impersonalization, some important strategies were lost over the course of the ME period, such as object fronting (topicalization and information-rearranging, see Speyer 2010) and *man* (impersonalization, Los 2002). As a result, the reliance on the passive for expressing these functions is likely to have become greater.

This section is structured as follows. Section 6.2.1 aims to establish whether passives in general became more frequent in Early Modern and MBE, by investigating both long and short passives, using different means to measure their frequency. Section 6.2.2 investigates the function of the passive after ME, examining the increased importance of the passive as an information-rearranger as well as investigating the topicalization function of the passive. Section 6.2.3 concludes.

6.2.1 The frequency of short and long passives

The hypothesis that the passive acts as a response to the loss of V2 leads to two expectations about its development after the ME period: (i) an increase in all passives, following the loss of alternative strategies for its topicalization and impersonalization functions and an increased need for subjects; (ii) an increase in long passives, following the loss of other information-rearranging strategies.⁷⁹ It is important to note that the development described here—an increase in the use of passives—is not a grammatical change as such, which means that it is not expected to develop in the same way as previously described syntactic changes (cf. Kroch's (1989) Constant Rate Hypothesis, which states that the rate of a change is similar in all contexts). Most importantly, there are no clearly-defined contexts in which passives are expected to replace actives, as in the case described by Kroch (1989), the spread of periphrastic *do*. In the case of the passive, the relevant construction is already (reasonably frequently) used in the language and there is no clear expected endpoint: it is not a matter of the passive being used in all possible contexts. Mostly, what we are interested in is the balance between active and passive. I show in this section that while there are some methodological challenges and there is considerable variation in the frequency of passives between individual texts, there is sufficient evidence to state that the passive becomes more frequent between the early sixteenth century and the early twentieth century.

6.2.1.1 Frequency of the passive

This first study investigates the frequency of the passive construction in EModE and MBE, in order to establish whether there is an increase in the use of the passive after the ME period. The studies in this section include both long and short passives, as well as all types of passives—direct, recipient, prepositional, ECM—because these are all likely to be affected by the development towards the need for more subjects and the importance of the passives as a topicalization and impersonalization construction. I present data on the frequency of passives using two different measures to compare the data from PPCEME and PPCMBE to available data for Early English, based on the frequency of passives with respect to actives (Seoane 2000, 2006), and data for PDE, based on frequency of passives with respect to the number of words (Evert 2006).

Method and selection

Passive clauses can be more easily selected in the syntactically annotated corpora for EModE and MBE (PPCEME and PPCMBE, respectively) than in the YCOE (see Section 4.2.2.1) because of the use of a separate label for passive participles. As in the YCOE, it is not possible to distinguish between verbal and adjectival passives, which means

⁷⁹ This section elaborates on material that was previously presented in Dreschler (2014).

that the selection is only based on form: all clauses with a passive participle and a form of *to be* as main clause constituents were selected. As well as being straightforward, this has the benefit of being close in selection to Seoane (2000), who also does not distinguish between verbal and adjectival use of the participle but (manually) selects sentences with a form of *to be* and a participle. The basic query for passives is given below.

<i>Query 6.1</i> Passive clauses in the PPCEME and PPCMBE	
a.	IP* idoms *SBJ* and IP* idoms *VAN*
b.	<i>either:</i> IP* idoms *BED* *BEP*
c.	<i>or:</i> IP* idoms *DOD* *DOP* *HVD* *HVP* *MD* *VBD* *VBP* *BED* *BEP* and IP* idoms BE BEN BAG

Line 6.1a selects all main and subordinate clauses (IP*) with a subject (*SBJ*) and a passive participle (*VAN*). Line 6.1b selects clauses that have a finite form of *to be*, while Line 6.1c selects clauses with a non-finite form of *to be* in addition to a finite verb.⁸⁰ In contrast to studies in the previous chapter, empty (e.g. moved and ellipled) subjects are also included, because the position of the subject is not relevant for the present investigation.⁸¹ The requirement for a finite verb in the clause ensures the exclusion of fragments and reduced relative clauses, which do not have an overt subject.

The selection of active transitive sentences was restricted to clauses with NP objects (coded as either NP-OB1 or NP-OB2). While it can be argued that other types of active clauses should also be taken into account in a comparison to the passive—most notably the types of actives that can be transformed into a prepositional passive and theme passives of ditransitives—there is unfortunately no straightforward way to select these clauses in the corpora: prepositional adjuncts are not distinguished from prepositional arguments, “for want of reliable criteria” (Santorini 2010). This means that the current selection may exclude some examples that are included by Seoane, who selected, by hand, all actives “for which a passive counterpart would be available” (2000:28). Query 6.2 was used to select the active clauses.

<i>Query 6.2</i> Active transitive clauses in the PPCEME and PPCMBE	
a.	IP* idoms *SBJ*
b.	and IP* idoms *DOD* *DOP* *HVD* *HVP* *MD* *VBD* *VBP* *BED* *BEP*
c.	and IP* idoms NP-OB*
d.	and IP* idoms !*VAN*

⁸⁰ Imperatives were excluded from the selection (for *be*, this would be BEI)

⁸¹ In *CorpusSearch*, this means that traces were not ignored.

Line 6.2a-c select all clauses with a subject, a finite verb and a direct object. Line 6.2d stipulates that this clause should not also dominate a passive participle. This extra condition was necessary to exclude clauses which have both a passive participle and an object, specifically passives of ditransitives (such as *It was told the knight*, see Section 6.3.2).

Results and discussion

Let us first examine how the queries described above compare to Seoane's method of selection and to what extent they confirm her results of an increase in the EMode period. The data presented below are based on Seoane's selection of genres and are, like her study, based only on the original Helsinki part of the PPCEME. Seoane selected three formal genres—*Statutes*, *Science* and *Sermons*—and three informal genres—*Private Letters*, *Drama* and *Fiction*. Table 6.1 first repeats part of Seoane's results and Table 6.2 shows the results based on Queries 6.1 and 6.2. The frequencies are given as percentages of the total number of passive and active clauses; the numbers in between brackets indicate the actual numbers.

Table 6.1 Relative frequency of passives with respect to active transitives, adapted from Seoane (2000: 28, her Table 3)

	E1	E2	E3
Total for these genres	21.4 (612,2236)	22.06 (722, 2550)	24.1 (922, 2893)

$\chi^2(2) = 7.81, p < 0.05$

Table 6.2 Relative frequency of passives with respect to active transitives in a selection of genres in the PPCEME

	E1	E2	E3
Statutes	46.4 (127,147)	55.7 (187,149)	64.6 (224,123)
Science	30.4 (119,273)	32.0 (175,372)	29.8 (146,344)
Sermons	20.4 (105,409)	19.2 (138,579)	18.3 (123,550)
Drama	11.0 (74,599)	8.8 (66,684)	10.6 (79,669)
Fiction	9.7 (64,595)	10.8 (81,667)	13.2 (94,616)
Private letters	14.0 (84,518)	14.2 (86,519)	16.3 (121,622)
Total	18.4 (573,2541)	19.8 (733,2970)	21.2 (787,2924)

For the last row only: $\chi^2(2) = 8.40, p < .05$

Table 6.2 shows that for all genres combined (the bottom row), there is a small but consistent increase over the entire selected period, from 18.4% in E1, to 19.8% in E2 and up to 21.2% in E3 (at roughly the same level of statistical significance as Seoane’s data). The table, then, confirms Seoane’s results of an increase in EmodE, although the percentages in Table 6.2 are lower overall, with the highest percentage in Table 6.2 roughly equal to the lowest percentage in Seoane’s results. In addition, the number of selected examples differs slightly per period—for instance, in E3, the total number of passives in Table 6.2 is 787 while it is 922 in Seoane’s data; in E2, the number of passives is roughly equal (722 from Seoane to 733 in Table 6.2) but the number of actives is quite different (2924 against 2550).

While the data confirm Seoane’s findings, the results in Table 6.2 also raise a number of questions with respect to the observation of an overall increase in these genres—some of these observations are less obvious in Seoane’s (2000) results because she does not give the breakdown of passives with respect to active transitives per genre. First, while the difference in frequency between the formal and informal genres is not unexpected—that was after all the reason for Seoane to select exactly these genres—Table 6.2 makes clear how large these differences are: in E2, for instance, the percentage of passives in *Statutes* is 55.7%, while in *Drama* it is only 8.8%. Despite these differences, the percentages overall per period are similar in all periods, at around twenty percent, which begs the question what this percentage actually says about the frequency of ‘the passive’ in these periods: can the genres be conflated at all to provide a meaningful percentage? Second, the increase in the total percentage per period is by no means a reflection of an overall development throughout all genres. Three genres do indeed show a consistent development: *Statutes* (48.7%, 59.1% and 66.4%), *Fiction* (7.9%, 10.8% and 11.9%) and *Private Letters* (14.0%, 14.2% and 16.3%), but others show a decrease (such as *Sermons*). Finally, the increase in E3 must be due to the high percentage for *Statutes*—already one of the two genres that have a much higher percentage of passives, as the column for ‘Total’ in Table 6.2 shows: 52.2%—in combination with the lack of an outlier below 10%, in contrast to E1 (*Fiction* at 9.7%) and E2 (*Drama* at 8.8%). Again, this raises the question whether the overall increase is a reflection of a development in the language as a whole during those periods, or whether it is only a development in these specific genres.

Despite these questions, the method of analysis seems to capture the distribution of actives and passives accurately enough to extend the selection of texts to the entire corpus—taking all genres and the additions to the PPCEME into account, as well as the PPCMBE. Table 6.3 below shows the results of Queries 6.1 and 6.2 for the entire PPCEME and PPCMBE.

Table 6.3 Relative frequency of passives with respect to active transitives in the PPCEME and PPCMBE

	E1	E2	E3	B1	B2	B3
Passive	6314	6572	5759	3531	4806	3336
Active transitive	22083	26480	20799	11644	12950	10441
Total	28397	33052	26558	15175	17756	13777
% Passive	22.2	19.9	21.7	23.3	27.1	24.2

For the entire table: χ^2 (5) 381.67, $p = 0$. Between each two subperiods: $p < .001$

Surprisingly, the EModE subperiods in Table 6.3 no longer show the increase found by Seoane and confirmed in Table 6.2; instead, there is a decrease from E1 to E2 (from 22.2% to 19.9%), which is followed by an increase up to 21.7% in E3, but this percentage is still lower than the percentage for E1. Unlike in the previous table, there is no consistent development throughout the periods: there is a decrease from E1 to E2, followed by an increase over three subperiods up to 27.1% in B2, and another drop in E3, down to 24.2%.⁸² We have already seen one possible explanation for this in the variation between genres: certain genres may show a considerably higher or lower percentage in a subperiod, influencing the overall percentage. In other words, the variation can be due to the characteristics of the corpus rather than representing a development in the language.

Let us turn to another approach, presented by Evert (2006), who measures passives in two PDE corpora⁸³ against the total number of words in the corpora. He uses the passive as an example in a discussion on the representativeness of corpora, showing that the number of passives varies greatly between texts, which makes it difficult to make a general statement about the frequency of the passive. In his two selected corpora, the frequency of passives⁸⁴ is roughly 12 per 1,000 words. The variation between texts becomes apparent when the overall frequency is compared to the frequency in Scientific texts, which is 17.3 per 1,000 words (2006: 180), a difference of 5.5 per 1,000 words. One very telling illustration of the variability in use of the passive is the frequency in books by Charles Dickens: the lowest number is 7.5 per 1,000 words for *A Christmas Carol*, while *Three Ghost Stories* has 11.0 passives per 1,000 words. Evert's discussion shows that the problems encountered for the PPCEME and PPCMBE are not only due to the challenges presented by historical corpora, but that the variability between genres and texts represents a problem for all corpora. In a strict interpretation, this suggests that it is only possible to consider

⁸² I considered main clauses separately because of the interest in V2, a main clause phenomenon. The development in main clauses is similar to the development in all clauses, but the percentages are 2 to 3 per cent lower.

⁸³ The *Brown* corpus (Francis and Kucera 1964) and the *British National Corpus* (BNC, Aston and Burnard 1998, written part only).

⁸⁴ Passive is defined as a form of *to be* followed by a past participle (179, fn3). Evert also gives the number of passives per number of verbal units but these are defined on the basis of his corpus so cannot be reproduced for the PPCEME and PPCMBE.

the frequency of passives for a single text, but we will consider a different approach below. In any case, this type of variation underlines the relevance of knowing the corpus and the fact that a corpus contains a balanced selection of genres may still not be representative or balanced with respect to use of passives.

Evert's data allow for a comparison of the frequency in the EModE and MBE corpora with his data for PDE. Table 6.4 shows the results of the frequency of passives in all clauses, main and subordinate, per 1,000 words.

Table 6.4 Frequency of passives per 1,000 words in the PPCEME and PPCMBE

	E1	E2	E3	B1	B2	B3
Passives	6314	6572	5759	3531	4806	3336
No. of words	567795	628463	541595	298764	368804	281327
Passives per 1000 words	11.120	10.457	10.633	11.819	13.031	11.858

For the entire table: $\chi^2(5) 173.21, p = 0$. Between each two subperiods: $p < .001$, except E2 and E3: $p = .36$

The numbers in Table 6.4 are comparable to Evert's result of 12 passives per 1,000 words: they range from 10.500 to 12.000, with the exception of the much higher number in B2 at 13.031. What is most important about this table, however, is that it shows the same decrease during the EmodE period that we saw in Table 6.3 above and which contradicts Seoane's results and the results in Table 6.2. Interestingly, the diachronic pattern throughout the subperiods is similar to the one found in Table 6.3: a decrease from E1 to E2, then an increase for each subperiod up to B2, and a final decrease from B2 to B3. An important conclusion, then, is that the variation is not due to one specific measure used, but that both selection measures capture this same development in the corpora.

We now have data for a development across the subperiods that is confirmed by both measures of frequency, but there is still the question of how this development in the percentages per subperiod should be interpreted. One way to investigate this is to single out specific genres—not to achieve a balanced overall frequency across genres, but rather to consider developments for separate genres over a number of periods. First, let us consider *Fiction* in MBE. Although this is a genre with low percentages of passives and considerable differences between authors, it is also a genre which is consistently present in all subperiods and which showed an increase in EModE (from 9.7% to 13.2%, see table 6.2). Table 6.5 presents the data for this genre only, based on frequency of passives with respect to actives.

Table 6.5 Relative frequency of passives with respect to active transitives in *Fiction* texts in the PPCEME and PPCMBE

	B1	B2	B3
Passive	46	122	75
Active transitive	378	735	433
Total	424	857	508
% Passive	10.8	14.2	14.8

Entire table: χ^2 (2) 3.61, $p = .16$. Between B1 and B2: $p = .09$. Between B2 and B3: $p = .79$

Table 6.5 shows an increase of passives, from 10.8% in B1 to 14.8% in B3, but this does not represent a consistent increase building on the EModE period: the percentage for E3 in Table 6.2 was 13.2%.⁸⁵ However, because the number of texts is small, these percentages alone do not provide sufficient proof for an increase; the differences are not statistically significant. Another genre that lends itself particularly well for a comparison over time is that of bible translations, because in addition to being present in many (though not all) subperiods, the texts selected for each subperiod in these corpora are based on the same sections of the Bible.

Table 6.6 Relative frequency of passives with respect to active transitives in *Bible* texts in the PPCEME and PPCMBE

	E1	E2	B1	B2
Passive	216	256	54	75
Active transitive	2076	2103	559	704
Total	2292	2359	613	779
% Passive	9.4	10.9	8.8	9.6

χ^2 (3) 3.83 $p = .28$. Between E1 and E2: $p = .11$, Between E2 and B1: $p = .14$, Between B1 and B2: $p = .60$

These results do not show an increase over the course of these periods, but the reliability of this is questioned by the fact that the differences between the highest and lowest percentages are quite small (8.8% and 10.9%); as in the previous table, the differences are not statistically significant. In conclusion, while selecting specific genres to investigate a diachronic development means that certain problems are avoided (such as differences between genres), it also means that the numbers are lower and it is difficult to determine whether the variation that is found represents a diachronic development or whether it is due to the characteristics of a single text (or translation).

The second solution to the problem of the variation between texts and genres is to make the selection as general as possible, which abstracts away from the

⁸⁵ Interestingly, the main clauses in Table 6.5 show a different development, with roughly similar percentages in B1 and B2 (13.1% and 13.3%, respectively) and a higher percentage in B2 (16.0%).

individual differences between texts; this may tell us something about the general development in the language. Table 6.7 shows the results of the queries for actives and passives (cf. Table 6.3 above) for each corpus, and Table 6.8 shows the frequencies for passives with respect to the number of words (cf. Table 6.4).

Table 6.7 Relative frequency of passives with respect to active transitives in the entire PPCEME and PPCMBE

	EModE	MBE
Passive	18645	11673
Active transitive	69362	35035
% Passive	21.2	25.0

χ^2 (1) 253.39, $p = 0$

Table 6.8 Frequency of passives per 1,000 words in the entire PPCEME and PPCMBE

	EModE	MBE
Passives	18645	11673
No of words	1737853	948895
Per 1000 words	10.729	12.302

χ^2 (1) 133.01, $p = 0$

Both tables show a quite substantial, and statistically significant, increase from EModE to MBE: from 21.2% to 25.0% in Table 6.7 and from 10.729 to 12.302 in Table 6.8. While this method of considering the diachronic development does not solve all problems that have been discussed in this section, especially the influence of genre and the representativeness of the (or any) corpora, the fact that there is such a considerable difference in percentages between EmodE and MBE suggests that it is likely that at least some increase takes place in the transition between the two corpora. What we cannot say on the basis of the data presented in this section is when exactly in the subperiods the major part of this change takes place. This is not surprising because the change that we expect as the effect of the loss of V2 is not a straightforward grammatical change but a change in frequency for which the starting point and endpoint, as well as the rate of change, are not well-defined; it is unlikely that the change takes place within the course of half a century.

In conclusion, there is evidence that the passive becomes more frequent after the ME period, although the increase can only be shown on a general level and cannot be shown to develop consistently throughout the centuries. The increase in the EmodE period that was earlier reported by Seoane (2000) for a selection of texts was confirmed, but could not be extended to the entire EModE and the MBE corpora. Several reasons were discussed to explain this effect, such as the influence of genre and variation between individual texts belonging to the same genre. The data for the two entire corpora, abstracting away from details of individual texts, show that overall there is an increase from EModE to MBE (i.e. around 1700), both of passives with respect to actives and of passives per 1,000 words. This means that

while it is not possible, on the basis of these data, to determine the dates for the development in more detail throughout the subperiods, the overall change shows an impact in the texts at least two centuries after the loss of V2, which coincides with the development in the subject as described in the previous chapter.

6.2.1.2 Frequency of long passives

While the need for subject strategies is relevant for all passives, the loss of alternative information-rearranging constructions such as object fronting and PP preposing specifically applies to long passives. Recall from Chapter 2 that one of the crucial functions of the passive, as shown in a corpus study by Birner (1998), is to create a given-before-new order of information. Example (1) illustrates this principle.

- (1) *Before the preparations were finished , Constantius received the intelligence of the tyrant's death, and it was considered as a sure presage of the approaching victory. The servants of Carausius imitated the example of treason , which he had given.*

He was murdered by his first minister Allectus

and the assassin succeeded to his power and to his danger. But he possessed not equal abilities either to exercise the one, or to repel the other.

(b2, gibbon-1776,1,365.165-170)

In the passive sentence, the subject *He* represents OLD information, referring back to the subject of the previous sentences. The *by*-phrase *by his first minister Allectus*, in contrast, is ANCHORED: Allectus himself has not been mentioned before and is therefore new to the discourse—although his presence is mediated by the introduction *his first minister*, which includes a referential pronoun. Had this sentence been an active, the unmarked order of information would have been violated: the subject would have been ANCHORED and the object OLD. The corpus study in Section 4.2 showed that in OE the passive was already used to rearrange information, but that this function was less pronounced because the passive was not yet completely fixed as a grammatical construction (for instance, the *by*-phrase had not yet grammaticalized) and because there were alternatives to the rearranging function, such as object fronting. On the assumption that creating given-before-new order remains similarly relevant over time (cf. Lambrecht 1994: 29), we expect that when these other means are lost after the ME period, the long passive becomes more frequent.

The selection of long passives in the corpus, like the selection of passives themselves, has become somewhat more straightforward in the PPCEME and PPCMBE corpora because of the establishment of *by* as the standard preposition for the agent phrase (e.g. Peitsara 1992). Query 6.3 presents the queries that were used to select passive main clauses with a prepositional phrase headed by *by*.

<p>Query 6.3 Passive clauses with <i>by</i>-PPs</p> <p>a. IP* idoms *SBJ* and IP* idoms PP* and PP* idoms P* and P* idoms By by</p> <p>b. <i>either</i>: IP* idoms *BED* *BEP* and IP* idoms *VAN*</p> <p>c. <i>or</i>: IP* idoms *DOD* *DOP* *HVD* *HVP* *MD* *VBD* *VBP* *BED* *BEP* and IP* idoms BE BAG BEN and IP* idoms *VAN*</p>

Line 6.3a selects all clauses with a subject, and a PP with *by* as the preposition. Lines 6.3b and 6.3c select clauses containing a passive participle and a form of *to be*, either finite (as in 6.3b) or non-finite, following a finite verb (line 6.3c). However, we have seen in Chapter 2 that *by*-phrases do not always encode agency, but can express different semantic roles (e.g. manner or means). Indeed, in addition to agentive *by*-phrases as in (2), Query 6.3 also selects a range of other *by*-phrases, as illustrated in (3)-(5).

(2) *I **was always treated** by him with the utmost Affection and Tenderness.*
(b1, anon-1711,6.44)

(3) *By what we think trifles, the worst habits **are often produced**; so that the mere neglect of external forms, is sometimes the cause of great disorder.*
(b1, barclay-1743,34.243)

(4) *for by the woods and rocks which softened at his song, **were meant** the first rude sons of men, whom his persuasive eloquence refined and melted into love.*
(b1, barclay-1743,17.66)

(5) *All the Modern Languages, as he says, **are learn'd** to a great Perfection only by Conversation;*
(b1, anon-1711,7.67)

While in (2), the *by*-phrase *by him* is clearly the agent for the verb *treat*, this is not the case for the other examples: *By what we think trifles* in (3) can arguably still be interpreted as an agent-like—even though a corresponding active would sound odd (*Trifles produce the worst habits*), possibly because the *by*-phrase is not animate. In examples (4) and (5), however, the *by*-phrase encodes some sort of means (*by Conversation*) or a general theme (*by the woods and rocks*). While, technically speaking, these examples may not represent true long passives, the point has already been made before that in view of the information-structural aims, the exact semantic role of the *by*-phrase is not as crucial as it would be for a different type of analysis; the information ordering principles still apply.

That being said, it might prove useful to consider a way to increase the likelihood that a selected *by*-phrase is at least closely associated with the passive participle in the sentence. One option is to only select those clauses in which the PP immediately follows the passive participle. The proximity between participle and *by*-phrase in a clause, as in (2), can indicate a close thematic relation, although there is no guarantee that this is indeed the case. The clauses were selected using the additional query line in Query 6.4 below, which was combined with Query 6.3 (specifically, line 6.3a). The query in 6.4 requires that the passive participle immediately precedes the prepositional phrase. Table 6.9 shows the results of both Query 6.3 and the combination of Query 6.3 and 6.4.

Query 6.4 Passive clauses with <i>by</i> -PPs immediately following the passive participle
a. *VAN* precedes PP*

Table 6.9 PPs headed by *by* in passive clauses in the PPCEME and PPCMBE

	E1	E2	E3	Total	B1	B2	B3	Total
Passives	6265	6540	5714	18519	3526	4789	3333	11648
VAN(...)by-PP	708	802	879	2389	567	720	454	1741
% VAN(...)by-PP	11.3	12.3	15.4	12.9	16.1	15.0	13.6	14.9
VAN-by-PP	331	345	410	1086	277	336	211	824
% VAN-by-PP	5.3	5.3	7.3	5.9	7.9	7.0	6.3	7.1

For the total per corpus only: VAN(...)by-PP χ^2 (1) 25.35, $p < .001$; VAN-by-PP χ^2 (1) 17.65 $p < .001$

The percentage of long passives as a proportion of all passives is quite low: ranging from 11.3% to 16.1% for VAN(...)PP, and from 5.3% to 7.9% for VAN-PP. These percentages are quite a bit lower than those summarized by Toyota (2008) for PDE, who estimated this to be around 20 to 30 per cent. They are more comparable to Seoane's (2000: 28) percentage of 15.4% for the entire EModE. What these data show is an increase in the proportion of long passives from the Early Modern corpus to the Modern corpus, both in the VAN(...)PP and the VAN-PP clauses: from 12.8% to 14.9% and 5.8% to 7.1%, respectively. Like in the data for the overall frequency of passives, however, this increase only becomes visible when comparing the two entire corpora and does not represent a consistent increase throughout the subperiods of these corpora: for both types of clauses, there is an increase up to the B1 period, which is followed by a decrease over two periods. Unfortunately, then, it is not possible to provide a more specific time for the change.

Finally, this is also the period in which the position of the *by*-phrase becomes more stable. While in E1 roughly a quarter (25.6%) of all *by*-phrases occur before the passive participle, this percentage declines steadily throughout the subperiods to

4.4% in B3.⁸⁶ This is in line with the decrease in PPs occurring in initial position, as described in Chapter 5; while *by*-phrases can be used for local anchoring in the case that the *by*-phrase is more given than the subject (see also Chapter 4, Section 4.2.2.2), this option ceases to be available.

In conclusion, there is evidence that in addition to the overall increase in the use of passives from EModE to MBE, there is also an increase in the use of long passives, while the position of the *by*-phrase becomes more fixed, following the passive participle. The increase is reminiscent of the changes in frequencies of passives that we saw before, with an overall increase from EModE and MBE, but no consistent increase throughout all subperiods. Like the increase in the use of all passives, the change takes place quite some time after the loss of V2, which makes it close to the developments towards a subject-initial language as described in the previous chapter.

6.2.1.3 Conclusion

The two different ways of measuring the frequency of all passives in the language—with respect to active transitive sentences and to the number of words—showed a similar development: no consistent development throughout the subperiods, but an increase in the use of passives from the EModE period to the MBE period. These two seemingly contradictory findings are not surprising, and for two reasons: first, there is considerable variation between genres and even between individual texts belonging to the same genre, which makes it difficult to assess overall changes; and second, these changes in frequency and use of a construction are expected to show a different rate of change—slower and less constant—from situations in which one grammatical option is replaced by another. The second study in this section focused specifically on long passives, because they are expected to become more important as one of the few remaining means for information-rearranging. The findings were comparable to those for the overall frequency of the passive: a general increase over all periods, but no consistent development from one subperiod to the next. In conclusion, while the data do not show in detail what the turning point of the changes in frequency of the passive—it does not seem to be a single consistent development—there are clear indications for an overall increase over the course of the centuries following the loss of V2.

6.2.2 The relevance of the passive as an information re-arranger after ME

The second aspect that is relevant for the question of an increased importance for the passive in the centuries after the loss of V2 is a functional change during these centuries. Birner (1996) and Seoane (2000) have shown on the basis of corpus studies that the passive has an information-rearranging function in both PDE and

⁸⁶ The following query line was added to Query 6.3: PP* precedes *VAN*

EModE, and the data in Chapter 4 showed that this function was already present in OE as well, although the extent to which the passive functioned as an information-rearranging device was more limited, both because of the functional competition with object fronting and PP preposing, and because of the unstable syntactic status of the passive (for instance, the lack of a standardized way of expressing the agent). Considering that the rearranging function has already been established, this section investigates two separate possible changes in the information-rearranging function of the passive from OE to the end of the EModE period: section 6.2.2.1 investigates the subject-creating function of the passive and section 6.2.2.2 proposes a separate function for the newness of the agent, of topic introduction.

6.2.2.1 The subject-creating or discourse-linking function of the passive

The information-structural function of the passive is not only a matter of the relative information status of the arguments, but Seoane's (2006) and Los's (2009) scenarios for the connection between the loss of V2 and the passive also build on the notion that passives provide a means for the language to create subjects. Incorporated in this proposal is the possibility that subject-creation is not only a syntactic mechanism, but also an information-structural device, because by placing the old(er) patient in subject position, the passive creates a given, and therefore unmarked, subject. This leads to the question whether, as the subject becomes more frequent as the initial element and gains importance because of loss of alternatives—not only rearranging alternatives such as object fronting, but also impersonalization strategies such as *man* (Los 2002)—this subject-creating function becomes visible in results for discourse-linking properties of the passive subjects. The data in Section 4.2.2.3 showed that in OE there were no indications for a difference in behaviour between passive and active subjects: their information status was comparable, although passive clauses were less often subject-initial than active clauses. This section provides a follow-up to those data and aims to establish whether the discourse-linking properties of passive subjects—specifically information status and clause-initial position—are different from active subjects and whether they have changed over the course of the centuries.

Data on the first discourse-linking property, the information status, is available from Seoane's (2000) data on long passives in EModE. Table 6.10 repeats her results (cf. Chapter 2, Section 2.3.2), but now only gives the information status of the subject, without the comparison to the information status of the agent.

Table 6.10 Information status of the subject in long passives in EModE (adapted from Seoane 2000: 29, her Table 5)⁸⁷

Information status of subject	n	%
Given	238	68.2
New	111	31.8
Total	349	100.0

The table confirms the general tendency for passive subjects to be old, with 68.2% given, but at the same time, the percentage of old subjects is lower than the percentage reported in Chapter 4 for *Orosius* and *Lives of Saints*, where we found that 56 out of 63 (88.9%) of the passive subjects were OLD (Table 4.16). The high percentage of new subjects (31.8%) compared to OE (11.1%) is particularly striking. One possible explanation emerges when we consider separately the genres that Seoane investigated: while the data for OE were based on two specific narrative genres, Seoane included a wider range of genres in her investigation. Some of these genres have a higher-than-average percentage of new subjects: *statutes* (32.7%), *private letters* (54.2%) and *sermons* (34.2%) and these are genres that are not straightforwardly narrative, which may explain their different behaviour to some extent (2000: 29, her Table 6). At the same time, we need to take into account Birner & Ward's observation that there is no absolute condition on the subject to be given, as well as the fact that passive sentences, like any sentence, can have an event-reporting character (Lambrecht 1994: 124), meaning that the entire event that is related is new and so the subject itself may also be new.

Having considered the subjects of passives separately, let us now turn to a comparison of active and passive subjects, following the reasoning that, given the specific subject-creating function of the passive, passive subjects are placed in initial position because of the need to place old unmarked arguments in subject position. As in section 4.2.2.3, I attempted to find a way to approximate the information status of subjects in the corpus on the basis of queries only; for given subjects, I selected personal pronouns, and for new subjects I selected NPs with an indefinite article. Let us first consider the proportion of pronominal subjects in active and passive clauses in EModE and MBE. Active and passive clauses were selected with the same queries as in Query 6.1 and 6.2, without the requirement for the active clauses to have an object. Additional queries were used to select the different types of subjects, and no selection was made on the position for the subject.⁸⁸ Table 6.11 presents the results.

⁸⁷ Seoane (2006: 378) presents an extension of here database, resulting in 72.5% given subjects (n=319) and 27.5% new subjects (121).

⁸⁸ IP* idoms NP-SBJ* and NP-SBJ* idomsonly PRO*

Table 6.11 Percentage of pronominal subjects in active and passive clauses in the PPCEME and PPCMBE

		PPCEME	PPCMBE
Active	Pronominal subjects	29139	16417
	Total active	59118	35456
	% pronominal	49.3	46.3
Passive	Pronominal subjects	2192	1512
	Total passive	5772	4545
	% pronominal	38.0	33.3

Active clauses, between PPCEME and PPCMBE: $\chi^2(1) = 79.21$, $p = 0$; Passive clauses: $\chi^2(1) = 24.50$, $p < .001$
 Between active and passive clauses in PPCEME: $\chi^2(1) = 269.53$, $p = 0$; in PPCMBE: $\chi^2(1) = 276.78$, $p = 0$

This table presents two unexpected results. First, the percentages of pronominal subjects in both active and passive clauses are considerably higher than in the OE data: they range from 33.3% to 49.3%, against figures around 21% for *Orosius* and *Lives of Saints*, and 31% for the entire YCOE (see Section 4.2.2.3). It is not immediately clear, on the basis of only these numbers and no analysis of individual examples, why such a difference might exist, although it is clear that the discourse organization, as well as the corpus annotation, may be different for OE than EModE and MBE. Second, the percentages of pronominal subjects are higher for actives (49.3% and 46.3%) than for passives (38.0% and 33.3%), which means there is no indication that passive subjects are more often given than active subjects. The percentages of pronominal subjects decrease in the transition from EModE to MBE, both for active and passive clauses. Again, these data are not detailed enough to provide an explanation as to why this development takes place; we also have to take into account that pronominal subjects provide only an approximation of given subjects.

Turning from old to new subjects, Table 6.12 presents the results of a query selecting indefinite subjects, again based on Queries 6.1 and 6.2, but now with an extra requirement that the subject should contain an indefinite article.⁸⁹ The percentages of indefinite subjects in Table 6.12 are low for both active and passive clauses, but considerably higher for passive subjects (1.0% and 1.6% for active subjects against 3.0% and 5.3% for passive subjects). Both active and passive clauses show an increase from EModE to MBE. This points in the same direction as the observation that the number of given subjects decreases, but it is clear that these numbers only scratch the surface. Again, however, these data show the reverse of what we would expect if passive subjects are placed in subject position especially, and increasingly so, because of their old information status.

⁸⁹ NP-SBJ* idoms D* and D* idoms a|an|A|An

Table 6.12 Percentage of indefinite subjects in active and passive clauses in the PPCEME and PPCMBE

		PPCEME	PPCMBE
Active	Indefinite subjects	580	553
	Total active	59118	35456
	% pronominal	1.0	1.6
Passive	Indefinite subjects	171	242
	Total active	5772	4545
	% pronominal	3.0	5.3

Active clauses, between PPCEME and PPCMBE: $\chi^2(1) = 70.07$, $p = 0$; Passive clauses: $\chi^2(1) = 36.91$, $p = 0$. Between active and passive clauses in PPCEME: $\chi^2(1) = 180.49$, $p = 0$; in PPCMBE: $\chi^2(1) = 279.70$, $p = 0$.

As a final approximation of the discourse-linking function of passive and active subjects we can consider the position of the subjects, under the assumption that a subject is more likely to play a discourse-linking role—especially repairing for loss of local anchoring—when it is placed clause-initially. Query 6.5 below was used to select subjects in clause-initial position in active and passive clauses, disregarding conjunctions and some technical nodes (cf. Section 5.3.1).

<i>Query 6.5</i> Subject-initial conjunct and non-conjunct clauses in the PPCEME and PPCMBE
a. IP* idomsfirst NP-SBJ*
The following nodes were ignored: ** QTP META LATIN FW CODE LS CONJ*

Table 6.13 Clause-initial subjects in active and passive clauses

	E1	E2	E3	B1	B2	B3
Active clauses	39186	46762	38149	21494	24924	20842
Subject-initial	26648	31446	27335	15691	18458	16021
% Subject-initial	68.0	67.2	71.7	73.0	74.1	76.9
Passive clauses	4956	5217	4413	2816	3961	2808
Subject-initial	3325	3540	3296	2114	2998	2187
% Subject-initial	67.1	67.9	74.7	75.1	75.7	77.9

In the first two subperiods, the percentages of subject-initial clauses for actives and passives are very similar, both between 67% and 68%, but from E3 onwards, the percentages diverge. Both actives and passives show an increase, reflecting the increase in subject-initial clauses that we saw in Section 5.3.1. What we can now see, however, is that this increase is slightly more pronounced for passive subjects, with somewhat higher percentages, although the differences are small.

In conclusion, the studies in this section have provided no evidence that passive subjects behave differently from active subjects with respect to some central

approximations for discourse-linking functions: old subjects, new subjects and the clause-initial position. There are no changes from the OE data presented in Chapter 4 that we are able to interpret as diachronic developments. However, this does not mean that the discourse-linking properties of passive subjects are not relevant; in fact, it is exactly the fact that passive subjects behave in a similar way to active subjects which makes them relevant as subject strategies: the fact that they behave like active subjects shows that they fit into the discourse-linking pattern that is characteristic for all subjects.

6.2.2.2 The passive as a topic introducer

The second change in the function of the passive after the OE period is that the newness of the agent may have a separate function: it can accommodate a function of introducing a new referent into the discourse.⁹⁰ The increasingly fixed position for the agent phrase after the past participle, towards the end of the clause, means that the agent phrase provides an optimal position to introduce referents into the discourse. It has been established for PDE that the majority of agents in the *by*-phrase are discourse-new (see e.g. Biber et al.1999: 942, who claim that “about 90%” of the agent phrases contain new information) and that they generally occur in a postverbal position—the favoured position for new information in PDE. Indeed, Siewierska (1984) describes the rearranging function of the passive not only as “enabling a given patient to function as the unmarked topic or theme”, but she also points out that it enables “a new agent [to function] as the information focus” (1984: 222). This section aims to establish whether the function of introducing a new referent into the discourse in focus position was already present in EModE.

Introducing a new referent has previously been used to describe the function of a range of constructions in PDE, two of which are given in (6).

- (6) Once there was *a wizard*. *He* was very wise, rich, and was married to a beautiful witch. They had two sons. The first was tall and brooding, he spent his days in the forest hunting snails, and his mother was afraid of him. The second was short and vivacious, a bit crazy but always game. Now *the wizard*, *he* lived in Africa.

(Lambrecht 1994: 177)

Lambrecht proposes that both the presentational construction *Once there was a wizard* and the left-dislocation *Now the wizard, he lived in Africa* have the function of (re-)introducing a referent into the discourse. Lambrecht defines this as ‘topic promotion’: these constructions promote the referent from a low position on Lambrecht’s Topic Acceptability Scale, roughly corresponding to new, to a high position, roughly corresponding to given (1994: 165, 176). This allows the referent to

⁹⁰ This is a shortened and adapted version of Dreschler (2010).

become the topic of the next sentence: when a referent has been introduced, and is active in the discourse, it has the most acceptable status to become the topic of the next sentence. Lambrecht makes a distinction between the two constructions in terms of the information status of the referent: presentational constructions prefer brand-new arguments (178), while left-dislocation occurs mostly with referents that have been mentioned somewhere before in the discourse or are in some way accessible to the hearer, but still need promotion in order to occur as a topic (183). In addition, Lambrecht specifically states (for the presentation construction) that while the topic (re-)introduction constructions make the referent available, this “often (but not always)” leads to predication in the subsequent discourse. This suggests that predication in the subsequent discourse is not a necessary diagnostic for topic introduction.

A different type of definition for topic introduction is given by Gregory and Michaelis (2001), who focus on the referent’s presence in the discourse rather than its information status. Gregory and Michaelis present a corpus study of the Switchboard Telephone Corpus (Godfrey, Holliman & McDaniel 1992), in which they aim to establish the difference in function between topicalization and left-dislocation. They define what they call “topic establishment” as a combination of no presence in the preceding discourse in the sentence before the specific construction ($t-1$) and presence in the discourse after the construction is used ($t+1$). In addition, the information status of the left-dislocated or topicalized element is measured, i.e. whether the referent is in some way accessible to the hearer, based on Gundel et al.’s (1993) categories. Their analysis of the 177 instances of left-dislocation in their corpus shows that: (i) most of the referents are given or at least inferable; (ii) most of the referents have not been mentioned before in the selected context; and (iii) most of the referents persist as a topic in the subsequent discourse. Crucially, the topic introduction function as described by Gregory & Michaelis is not based on the information status, but only on discourse behaviour.

Returning to the hypothesis that the passive in EModE functions as a topic introducer, this means that the *by*-phrase agents of the long passives in the corpus are expected to show similar behaviour with regard to ‘givenness’, ‘anaphoricity’ and ‘persistence’ to the left-dislocation constructions in Gregory & Michaelis’ corpus. Based on Lambrecht’s discussion, an important question is whether the passive functions as a topic introducer of new elements or whether it makes referents that are already accessible to the hearer more acceptable to become the topic of the next sentence.

Method: corpus and selection

The corpus for both studies consisted of a selection of texts from the PPCME2 and PPCEME (the Helsinki part). Note that I have used the text files and not the annotated files. Only long passives with an animate agent in the *by*-phrase were selected, in order to avoid confusion with other uses of *by*, e.g. for expressing means. Table 6.14 presents an overview of the number of texts for each period that

is distinguished in the *Penn-Helsinki* corpora and the number of passives that were selected.

Table 6.14 Texts used in the corpus studies

Period	No. of texts	No. of words	Long passives	No. of long passives per 1,000 words
M4	12	260,116	56	0.215
E1	27	191,737	75	0.391
E2	28	191,074	138	0.722
E3	26	171,885	138	0.803

The long passives were analysed using three aspects of topic promotion: information status, anaphoricity, and persistence. These aspects were analysed in roughly the same way as in the other studies in this thesis (cf. Section 4.2.2.1), with two important differences: the context used in this study was 60 words instead of 5 main clauses; and for anaphoricity and persistence, I only used a binary distinction between “prior/subsequent mention” or “no prior/subsequent mention”. Table 6.15 repeats, once again, the information status categories.⁹¹

Table 6.15 (=Table 4.2) Annotation scheme for information status

Category	Categories from other models included in each category	Binary distinction
OLD	Prince’s (1981) <i>Textually Evoked</i> , Komen et al.’s (2014) <i>Identity</i> , Pintzuk & Taylor’s (2011, 2014) <i>Previously mentioned</i> .	‘Given’
ACCESSIBLE	Prince’s <i>Unused</i> and <i>Situationally Evoked</i> , Komen et al.’s <i>Assumed</i> , Taylor & Pintzuk’s <i>Shared/Cultural knowledge</i> , <i>Situationally Evoked</i> , Haug et al.’s (2014) <i>Generic</i> , Birner’s (2006) <i>Elaborating inferables</i>	
ANCHORED	Birner’s <i>Bridging inferables</i> , Prince’s <i>Brand-new Anchored</i>	‘New’
NEW	Prince’s <i>Brand-New Unanchored</i> , Taylor & Pintzuk’s and Haug et al.’s <i>Specific, Non-specific</i> and <i>Short-term referents</i> , Komen et al.’s <i>Inert</i>	

⁹¹ Note that the categories used are different from the original data as presented in Dreschler (2010), where, roughly speaking, the ACCESSIBLE and ANCHORED categories as shown in Table 6.16 were conflated into a category called ‘mediated’.

Results

Let us first investigate to what extent the long passives in the corpus conform to the definition of topic introduction as possible promotion on the givenness hierarchy (and with that on Lambrecht's Topic Acceptability Scale); in other words, a combination of the givenness and persistence measures. Table 6.16 shows the results of combining these two factors for the passive in M4 and EModE.

Table 6.16 Information status and persistence scores of *by*-phrases in M4 and EModE

	No persistence		Persistence		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
OLD	71	17.5	52	12.8	123	30.3
ACCESSIBLE	55	13.6	21	5.2	76	18.7
ANCHORED	84	20.7	42	10.3	126	31.0
NEW	35	8.6	46	11.3	81	20.0
<i>Total</i>	245	60.3	161	39.7	406	100.0

The table shows that there is an even distribution among the information status categories: OLD and ANCHORED have a percentage of roughly 30%, and ACCESSIBLE and NEW have a percentage of roughly 20%. Compared to Gregory & Michaelis's study, the category of new referents is much higher: only 6% of their referents belong to Gundel et al.'s (1993) two lowest givenness scales (*type identifiable* and *referential*). Yet this percentage is not as high as in Biber et al.'s (1999) percentages for PDE. The category that corresponds to the introduction and continued predication of a discourse-new referent is *new+persistence*. The table shows that only a small number of examples fall into this category: 11.3% (46 out of 406) of all passives in the corpus. One example is given in (7).

- (7) *for I may nat be delyverde of thys swerde but by a knight, and he muste be a passynge good man of hys hondys and of hys dedis, and withoute velony other trechory and withoute treson.*
(cmmalory,45.1497-8)

In example (7), a 'damsel' talks about a sword (old information) and then introduces a knight, who has not been mentioned before and is therefore new information. In the following sentence this newly-introduced knight is picked up as the speaker adds requirements for him. Because the number of examples is small, it seems that the strictest definition of topic introduction does not directly apply to the passive. As we have seen in both Gregory & Michaelis's and Lambrecht's discussions, however, the condition on persistence is not absolute: a topic introduction device makes referents available but it does not always need to be picked up; something which was especially relevant in Gregory & Michaelis's corpus, which consists of conversations—less predictable than planned text. Table 6.16 shows that 20.0% of the referents are new, which means that about one-fifth of the examples would still

meet the criteria for topic introduction. This is perhaps still not a large proportion, but in contrast to the rearranging function of the long passive, we do not expect the majority of long passives to perform this function.

Let us now consider the second definition of topic introduction, where topic introduction is understood as the introduction of a referent with any information status to the discourse, i.e. a referent which is not mentioned in the selected window. Topic introduction would then translate as a combination of no prior mention and persistence. Table 6.17 shows the results for M4 and EModE.

Table 6.17 Anaphoricity and persistence scores of *by*-phrases in M4 and EModE

	No persistence		Persistence		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Prior mention	36	8.9	35	8.6	71	17.5
No prior mention	209	51.5	126	31.0	335	82.5
<i>Total</i>	245	60.3	161	39.7	406	100.0

What we can now see is that the large majority of referents, 82.5%, are not mentioned in the immediately preceding context, even though they may be hearer-old.⁹² This is an even higher percentage than in Gregory and Michaelis's study: in their left-dislocation examples, a majority of 62% of the items had not been mentioned before. The category which now qualifies as topic introduction is considerably larger than under the previous definition, with 31.0% of the agent phrases having no prior mention combined with persistence, as illustrated in (8).

- (8) *air windoid in the north side of the market place, joining hard to the north west part of the bisshop's palace. This cumly peace of work **was made by Bisshop Bekington**, that myndid, yf he had lyvid lengger, to have buildid other xij. on the south side of the market steede, the which work if he h*
(cetrav1a.txt)

If we follow the same reasoning as before—that persistence is not a strict condition—all referents that have no prior mention can be described as cases of topic promotion, which amounts to 82.5% of all examples. What we have to take into account, however, is that the context selected for this study was quite small, which may affect the reliability of the anaphoricity results.

In conclusion, under both Lambrecht's and Gregory & Michaelis's definitions of topic introduction, there are a number of examples which meet the criteria for topic introduction, but more clearly so for topic promotion (31.0%) than for topic

⁹² Note that the largest category in this table, *no prior mention + no persistence*, are likely to represent typical cases of agent backgrounding: the referent occurs only once and is of little prominence to the discourse.

introduction (11.3%). This percentage becomes larger when actual subsequent mention is not taken as a necessary condition for (an attempt at) topic introduction: in that event, the percentage of new agents is 20.0% while the number of examples without a prior mention is even higher, at 82.5%. While it has to be taken into account that the window for anaphoricity and persistence is small, there seem to be enough indications that the passive is indeed used as a topic introduction device, but more so by way of promoting agents to become topics than by introducing new (information status) referents. This function of topic promotion, then, is an additional but complementary function to the subject-creating and information-rearranging functions of the long passive.

6.2.2.3 Conclusion

With Seoane (2000, 2006) already having established the information-rearranging function of the passive in EModE, this section focused on two related functions of the passive. The first was the subject-creating function of the language, in an attempt to see if this function—which we expect to become more important throughout the history of English—would become visible in the data. This did not prove to be the case: there were no indications for a different behaviour of passive subjects in this respect, both in comparison to active subjects and over time. However, I argued that this does not mean that the subject-creating function is not important but in fact, strengthens the idea that the passive makes use of the existing properties of subjects. The second aspect considered in this section was a possible separate function for the *by*-phrase in the long passive: when the passive became more standardized and the position of the *by*-phrase became fixed in end position in the clause, this means that the passive places a likely-to-be-new agent in focus position—a characteristic which the passive then shares with other constructions which have been described as topic introduction devices. An analysis of long passives in the last subperiod of the ME period and in the EModE period showed that there is indeed evidence that the passive was used as a topic introduction device in EModE. In conclusion, the data in this section show that the changes in the use of the passive are not only quantitative, but also functional.

6.2.3 Summary and conclusions

This section has investigated possible changes in the use of the passive after the ME period, focusing on both the frequency in use of the passive and the information-structural functions. The data in this section showed that despite some challenges in determining the frequency of passives and the development throughout the periods, there is enough evidence for an increase throughout the EModE and MBE periods. With respect to the information-structural function of the passive, the study of certain discourse-linking properties of the subjects of passives showed that there are no indications for a difference in behaviour between active and passive subjects. While this did not show evidence for a change in the function of passive subjects, I

argued that it underlined the function of the passive of creating acceptable, i.e. unmarked and given, subjects. The study of long passives in M4 and the EModE corpus showed that a proportion of long passives behaved as topic introducing devices. This topic introducing function of the passive only became relevant when the long passive itself had become an established construction: a standardized way of expressing the agent and a fixed position for the agent in the clause, namely towards the end, so in focus position.

In conclusion, what these studies have shown is that the centuries after the ME periods are crucial to the development of the passive, and that some of these developments are closely related to the overall development of the language towards more subject-initial clauses, as described in the previous chapter. While the data discussed in this section only scratch the surface of what is taking place, they provide an interesting outlook on the possible relation between the developments in the passive and the developments in the initial position in the clause.

Section 6.3 Development of the new passives after ME

This second section on passives after the ME period turns the attention to the three passives that were introduced in the fourteenth century: the prepositional passive, the recipient passive and the ECM passive. The studies in Chapter 3 showed that these passives were all in use by the end of the ME period, but their frequency was still quite low and they still seemed to be somewhat limited in syntactic productivity. The current section aims (i) to trace their development towards becoming a fixed part of the language, as they are in PDE; and (ii) to compare their development to the development of the passive as presented in the previous section, and to the loss of V2 as described in Chapter 5.

By the end of the ME period, the three passives are attested in the corpus but none of them seem as established a construction as they are in PDE. The use of these passives in the early periods is still quite different from what it is in PDE; for instance, both the theme passive and recipient passive are in use, while in PDE the recipient passive is the default option (in most dialects). Interestingly, while the introduction of these passives has been discussed in great detail in the literature, their subsequent development has been documented in far less detail, with only a few exceptions. For example, Visser (1963-73) discusses the development of the recipient and prepositional passive, and describes the development as a consistent increase from the introduction onwards. Similarly, Seoane (1999b) presents a corpus study of the recipient and theme passives in the EModE period and claims that this is the period in which they “consolidated”. For the ECM passive, Noël (2001, 2008) proposes that the ECM passive has grammaticalized into a fixed template, although this is based mostly on PDE material. Yet none of these studies presents a full account of these passives in a large corpus over a number of centuries.

Having established in Chapter 3 that the loss of V2 did not lead to the actual introduction of the new passives, the second part of the hypothesis that the

developments in the passive and the loss of V2 are connected is that the loss of V2 contributed to the further establishment of these passives in the language. Specifically, this would result in an increase in these new passives in the later periods, when the consequences of the loss of V2 kick in and there is an increased need for subject strategies and information-rearrangers. The previous section showed that there is a general increase in use of passives and long passives specifically in the Modern British period, and what this section aims to establish is whether the new passives are part of this development, or develop independently.

This section is structured as follows. Section 6.3.1 investigates the development of the prepositional passive, aiming to establish whether there is an increase in frequency as well as in types of verb+preposition collocations. Section 6.3.2 looks at the recipient passive, with the added question of how the competition between the recipient passive and the theme passive develops. Section 6.3.3, finally, turns to the ECM passive, and investigates the frequency of this passive, as well as the distribution of active and passive ECM. Section 6.3.4 concludes.

6.3.1 Prepositional passive

As we saw in Chapter 3, the prepositional passive is the first of the new passives to be introduced, and is also the passive which seems already most frequent and stable by the end of the ME period. In the ME data, the use of the prepositional passive is governed by several factors, such as recurring combinations of verbs and collocations and occurrence in relative clauses. In PDE, on the other hand, a different type of restriction seems to determine acceptability: the pragmatic constraints, as described by for instance Bolinger (1975) and Takami (1992). This section aims to trace the development of the prepositional passive between ME and PDE, both in terms of frequency and of use.

There is consensus in the literature that the prepositional passive was an established construction by the end of the fourteenth century (e.g. Denison 1993: 125, Mustanoja 1960: 440-441), but little data is available on the development of the prepositional passive after the fourteenth century, with two notable exceptions. Visser (1973: 2125, §1953) claims that the development of the prepositional passive after 1500 is a continuous increase over the centuries (“with undiminished speed” in the sixteenth century (1973: 2125)), with new verb+preposition combinations being attested in each century. He presents a long list of examples from a large selection of texts, but he does not show any frequencies per period. Seoane (1999b) also claims a “rapid spread” (135), although her data do not quite warrant such a description: based on a selection of texts from the Helsinki corpus (similar to the study of passives in Seoane 2000, 2006), she finds an increase in prepositional passives with respect to all passives from 0.9% in E1, through 3.3% in E2, to 2.3% in E3 (1999b: 134).

Seoane and Visser present different explanations for what they call the spread of the prepositional passive. Seoane (1999b: 135) attributes the rise to the increase in prepositional verbs in EModE, i.e. verbs taking prepositional complements, as well

as reanalysis of verb and preposition as one unit, although she does not explain the most important aspect of this reanalysis, namely why it should start to take place in the EModE period. Visser (1973: 2125) describes some examples in the sixteenth century already as having a “modern tinge” and calls them “forerunners” of the modern type *The bed was not slept in*. He also notes that in the sixteenth century most examples occur in colloquial texts. Finally, for the seventeenth-century data he mentions the advice from grammarians against preposition stranding, which is likely to affect the use of the prepositional passive as well—although his data do not suggest that the increase in frequency stops.

The following sections present the results of a corpus study of the prepositional passive in the PPCEME and PPCMBE, taking up the questions of frequency and possible limitations in use, both syntactic and pragmatic. The studies aim to establish whether (i) the prepositional passive continued to increase in frequency, as suggested by Visser (1973) and Seoane (199b); (ii) the minimal alternation contexts continued to make the prepositional passive possible; (iii) there are signs of the pragmatic constraint influencing the use of the prepositional passive in these periods.

Selection

Because of the similarities in coding, the same query that was used in Chapter 3 to select prepositional passives in the PPCME2 can be used for the PPCEME and PPCMBE.

Query 6.6 Prepositional passives in the PPCEME and PPCMBE
<p>a. IP* idoms *SBJ*</p> <p>b. IP* idoms PP* <i>and</i> PP* idoms P* <i>and</i> PP* idoms NP*<i>and</i> NP* idoms **</p> <p>c. IP* idoms *BED* *BEP* BE BAG BEN</p> <p>d. IP* idoms *VAN*</p>

Lines 6.6a to 6.6d select clauses with a subject, a stranded preposition, a form of *to be* and a passive participle. The query selects all types of stranded prepositions, making no distinction between the syntactic contexts for stranding, including examples such as (9) and (10).

- (9) *I undirstond by the Chaunceler touchynge the Duke of Gelders that he hath now late written a Lettre to the Kinge Catholyk off a more sobre style then he was wont to use, by which it may appere he makyth overture off treaty **whiche I thynke shal be herkynyd unto.*** (tunstall-e1-p1,3.1,273.31)

- (10) *And of thys towne was Joseph of Aramathia that awght the new Tumbe or Monymment **that our Savir Crist was buryed in.*** (torkingst-e1-h,25.37)

Only example (9) is a prepositional passive, i.e. only in that clause has the original complement of the stranded preposition (*overture off treaty*) passivized to become the subject of the verb (*shal be herkynyd unto*), followed by relativization of this subject. Example (10) may look like a prepositional passive, but the NP which is passivized to become the subject of the clause is *our Savir Crist*, the NP object of *bury*, not the complement of the preposition *in* (cf. *They buried him in the church* and *The church they buried him in*). The reason for stranding in this sentence is relativization, without passivization, of *the new Tumbe or Monymnt*. Only those clauses in which the subject of the passive clause represents the original complement of the preposition were included in the results.

Results and discussion

In order to determine whether the prepositional passive becomes more frequent after the ME period, its frequency is ideally considered against both the total number of passives and the number of active sentences with a prepositional object. Unfortunately, this second type of comparison is problematic because clauses with a verb and a prepositional phrase eligible to be passivized cannot be selected on the basis of form only (cf. also Section 6.2.1.1). Table 6.18 therefore presents the number of prepositional passives against the total number of passives in each subperiod, and against the total number of words in the corpus (cf. Evert 2006).

Table 6.18 Frequency of prepositional passives in the PPCEME and PPCMBE

Period	Prepositional passives	All passives	% Prepositional	No. of words	Per 1,000 words
E1	38	6282	0.60	567795	0.067
E2	99	6545	1.51	628463	0.158
E3	106	5741	1.85	541595	0.196
B1	58	3528	1.64	298764	0.194
B2	75	4797	1.56	368804	0.203
B3	75	3330	2.25	281327	0.267

Against all passives: **E1-E2: p < .001**; E2-E3: p = .16; E3-B1: p = .48; B1-B2: p = .78; **B2-B3: p < .05**

Per 1000 words: **E1-E2: p < .001**; E2-E3: p = .12; E3-B1: p = .96; B1-B2: p = .79; B2-B3: p = .10

Both ways of measuring the frequency show that the prepositional passive becomes more frequent throughout the EModE and MBE subperiods. Against the total number of words, there is a mostly continuous increase from 0.067 in E1 to 0.267 in B3, with the exception of a decrease of 0.002 from E3 to B1. The percentage of prepositional passives with respect to the total number of passives shows a somewhat different development: an increase from E1 to E2, from 0.60% to 1.51%, after which the percentage is variable, although it continues to be somewhat higher than 1.50%, until it reaches above 2 per cent, at 2.25%, in the final period. The

prepositional passive, then, continues to become more frequent in these periods, but the increase hardly makes it, in Visser's words, "one of the commonest usages in the language" (1973: 2125). This is not surprising when we consider the lexical restrictions for prepositional passives, i.e. the dependence on the availability of verbs taking prepositional complements.

There are some indications that the use of the prepositional passive is still limited in these periods. Most importantly, a considerable number of prepositional passives still occur in relative clauses, as in (11), suggesting that the acceptability of stranding in relative clauses continues to contribute to the acceptability of stranding in prepositional passives.

- (11) *Beware therefore lest that fall on you, which is **spoken of** in the Prophetes: Be holde ye despisers*
(tyndnew-e1-p2,13,40A.619)

In (11), *that* is the original complement of *of*, and it is both passivized and relativized. Table 6.19 shows the number of examples of prepositional passives in relative clauses or main clause *wh*-stranding environments with respect to the total number of prepositional passives in that period.

Table 6.19 Frequency of prepositional passives in relative and *wh*-movement clauses in the PPCEME and PPCMBE

	E1	E2	E3	B1	B2	B3
Prepositional passives	38	99	106	58	75	75
Prepositional passives in relative/<i>wh</i>-movement clauses	11	25	27	16	14	20
% Relative/<i>wh</i>-movement	28.9	25.3	25.5	27.6	18.7	26.7

$\chi^2(5) 1.39, p = .93$

Table 6.19 shows that just over a quarter (between 25.3% and 28.9%) of the prepositional passives occur in a relative clause or main clause *wh*-stranding environment, with the exception of the B2 period (18.7%). One way to interpret this is to say that stranding in prepositional passives still depends on, or is mediated by, the minimal alteration context of relative clauses where stranding is freer. However, the reliance on these stranding contexts can no longer account for all or even most of the prepositional passives.⁹³

⁹³ Interestingly, the number of relative clauses in which passives occur but where stranding is due to relativization only, so the option that was originally already possible but which may have been the way to a minimal alternation context, does show a decline over these periods.

While certain verb+preposition combinations occur more frequently than others, the sheer range of combinations that occur in the corpora shows that the prepositional passive is a fully productive pattern within the range of verbs taking prepositional complements or adjuncts. There is some evidence that there are limitations on productivity because of the dominance of some combinations that occur frequently in all periods, such as *sent for* in the EModE period: 3 of 38 examples in E1, 22 of 99 in E2 and 7 of 106 in E3. Moreover, some of these examples suggest that they are not real passives but some sort of fixed expression, since they are combined with other prepositions which do not logically combine with the meaning of *for*, as in (12) and (13).

(12) *A barber was **sent for** from the market towne hard by, who searcht his mouth,*
(armin-e2-h,13.154)

(13) *for I am **sent for** speedily to the Emperour of Trapezond, about affaires of great importance that highly concernes his royall person.*
(jotaylor-e2-p1,3,85.C1.252)

Other combinations that occur multiple times in (nearly) all periods are *spoken of*, *provided for* (only up to B1), *looked on* and *called on*. In most periods, there are also a number of verbs that occur much more frequently than all other verbs, such as *sent for* (22) and *spoke of* (10 examples) in E2, with the first verb after that occurring only 6 times; or, in E3, *disposed of* (11 examples), *looked on* (9 times) and *sent for* (7), with the next verb only having 4 examples. However, in some cases what may seem like a frequent occurrence of one specific combination is in fact due to many uses in the same text in the same context: for instance, in the text *weathers-1913*, 4 of the 6 prepositional passives are *dealt with*, which occur in a similar context. Moreover, the range of verb+preposition combinations that occur in the corpora provides support for the productivity of the prepositional passive. The occurrences in the entire corpus, 451, are divided among 249 verbs, meaning that any given verb+preposition combination on average occurs only approximately 1.8 times in the corpus (averages per subperiod range from 1.6 to 2.1). These numbers are higher in EModE than in MBE, meaning that the prepositional passives make use of a slightly wider range of verbs in MBE.

Another piece of evidence for the productivity of the prepositional passive are the combinations in which the relation between verb and preposition is less close than in fixed examples like *spoken of* and *provided for*. Some examples are given in (14)-(16).

(14) *and [the Matter] here was **sat upon** at Westminster in the Parliament-Chamber.*
(thoward2-e2-p1,1,89.216)

- (15) *you'd be **laught at** soundly,*
(middle-e2-p1,36.453)
- (16) *or in a large Pond to draw them to any certain place, that they may the better and with more hope be **fished for**, you are to throw into it in some certain place,*
(walton-e3-h,297.311)

Finally, there are some sentences with unexpected combinations for which it is clear that it is the context of 'affectedness' that makes the sentence acceptable.

- (17) *St. Winfreds Well is **built over** with stone on pillars like a tryumphall arch or tower on the gates of a Ch urch;*
(fiennes-e3-p2,180.288)
- (18) *The houses of the town are well built, but execrably furnished: bare, ghastly rooms, with an isolated chair or sofa sprinkled here and there, and looking as if they never **had been sat upon**, nor were meant to be.*
(reade-1863,222.536)

Example (17) represents the only example with *built over* in the two corpora, while *sat upon* occurs only twice. In both examples, there is a clear result described in the sentence, the buildings that stand on the well in (18) and the effect on the 'looker' that the chairs have never been used. These types of combinations, then, show that the pragmatics, frequently discussed for PDE, already widen the range of possible prepositional passives in earlier periods (cf. also Visser's (1973: 2125) remark on the sixteenth-century examples having a "modern tinge").

In conclusion, the prepositional passive becomes more frequent in the early EModE period, and remains at a fairly stable frequency after that until an increase towards the end of the MBE corpus (late nineteenth century). There are still some indications that stranding possibilities in relative clauses promote the use of the prepositional passive, while there are also some indications of lexical restrictions or fixed patterns. At the same time, however, the number of verb+preposition combinations shows that the prepositional passive is not limited to a specific range of verbs and it seems that the range of verbs becomes slightly wider over time. The increase in the prepositional passive is consistent throughout the subperiods and is consistent with the overall increase in the use of passives from EmodE to MBE.

6.3.2 Recipient passive

The recipient passive is the least established of the new passives by the end of the fourteenth century, with only a handful of examples attested in the corpus. What also makes the recipient passive special is the fact that other passives exist for the same type of active, and while the dative-fronted passive ('*Her was given the book*')

was no longer in use by the end of the ME period, the theme passive (*The book was given her*) continued to be available. At the same time, we have seen that in PDE, the recipient passive is the standard option and the theme passive is very rare (see Huddleston & Pullum 2002: 249). This section aims to establish whether the recipient passive shows an increase in use, like the prepositional passive, and investigates the balance between theme and recipient passives throughout the centuries.

As with the prepositional passive, only a limited amount of data is available for the recipient passive after its introduction. Visser (1973: 2144-2152, §§1968-1975) describes the development of the recipient passive after 1500 as another story of increase in frequency, of establishment, and of expansion to other verbs. This same claim is presented by Seoane (1999b), who, following Kisbye (1972), describes the recipient passive as “basically an EModE development” (136). This conclusion is, like her study for the prepositional passive, based on an examination of a part of the Helsinki corpus, and she finds that recipient passives come to make up a larger proportion of the passives in the language: 0.6% in E1, 2.07% in E2 and 1.4% in E3. However, both Visser and Seoane include clausal ditransitives in their studies, which we have seen in Chapter 3 behave and develop differently. When these passives are excluded from Seoane’s data, only four examples remain, all from E3, which can hardly be said to represent an “established” construction.

For the recipient passive, there is also a stylistic issue which may have slowed down the development (Visser 1973: 2149-2150, Denison 1993: 112). Visser cites various grammarians from the beginning of the twentieth century who advise against the use of the recipient passive. For instance, Sweet (1898) states that “we still ... try to evade such passive constructions as *she was given a watch*...because we still feel *she* and *he* are in the dative, not the accusative relation” (Sweet 1898: 118). Others simply state that it is wrong, calling it “loose Syntax” (Stokoe 1937: 78, *The Understanding of Syntax*, as quoted by Visser 1973: 2149). However, Visser also cites grammarians who accept the use of the recipient passive, some of whom were writing at an earlier date. Nevertheless, the existence of advice against the recipient passive and the availability of other options to passivize ditransitive verbs (most importantly with a prepositional recipient) are likely to have slowed down the development.

The following sections present the results of a corpus study of the recipient and theme passive in the PCEME and PPCMBE, which aims to determine (i) whether the recipient passive increases in frequency; (ii) whether it becomes a fully productive construction; and (iii) at what point the recipient passive replaces the theme passive as the most frequent passive for ditransitives with two NP objects.

Selection

The following queries, similar to those in Chapter 3, were used to select recipient and theme passives, based on the distinction made in the annotation of the PCEME

and PPCMBE between ‘first’ objects and ‘second’ objects, corresponding roughly to direct objects and indirect objects.

Query 6.7 Recipient and theme passives in the PPCEME and PPCMBE
a. IP* idoms *SBJ* and IP* idoms *BED* *BEP* BE BAG BEN and IP* idoms *VAN*
For recipient passives:
b. IP* idoms NP-OB1*
For theme passives:
c. IP* idoms NP-OB2*

The query for the theme passive is accurate in selecting theme passives (only 1 needed to be excluded), but the results for the query for recipient passives includes a range of clauses which are not recipient passives. Examples (19) and (20) illustrate two such cases.

(19) *But whan the kynge with his power was **entred** y=e= Countre he with his Fawtours fledde into the Mountaynes*
(fabyan-e1-h,170V.C2.114)

(20) *yea I am sorie, with all my harte, that they be **giuen** no more to riding, then they be:*
(asch-e1-p1,10R.186)

Example (19) seems to be a perfect rather than a passive and in example (20), *give* does not seem to have its usual meaning of ‘transfer’. Phrasal prepositional passives, as in (21) and (22), were also excluded, because the NP that is passivized may be part of a prepositional phrase rather than an independent argument.

(21) *and seems to have been known before Eugalenus had confounded most other diseases with the true scurvy; as it **is taken notice of** in the year 1624 by Sennertus, when Eugalenus 's writings, in all probability, might not have reached Norway.*
(lind-1753,264.150)

(22) *The mechanical truths brought to light by Archimedes, were, like his tomb, overgrown with the rank and unprofitable vegetation of later days, till they **were lost sight of**;*
(whewell-1837,23.219)

Results and discussion

Table 6.20 presents the frequency of recipient passives in the subperiods in the corpora. As before, the number of examples is given as a percentage of the total number of passives in the corpus, as well as against the total number of words.

Table 6.20 Frequency of recipient passives in the PPCEME and PPCMBE

Period	Recipient passives	All passives	% Recipient passives	No. of words	Recipient passives per 1,000 words
E1	5	6282	0.08	567795	0.009
E2	21	6545	0.32	628463	0.033
E3	24	5741	0.42	541595	0.044
B1	10	3528	0.28	298764	0.033
B2	10	4797	0.21	368804	0.027
B3	15	3330	0.45	281327	0.053

Against all passives: E1-E2: $p < .005$; E2-E3: $p = .38$; E3-B1: $p = .30$; B1-B2: $p = .49$; B2-B3: $p = .05$

Per 1000 words: E1-E2: $p < .005$; E2-E3: $p = .34$; E3-B1: $p = .45$; B1-B2: $p = .64$; B2-B3: $p = .25$

The table shows that the recipient passive continues to be a marginal construction for most of the EModE and MBE periods: as a proportion of all passives found in the corpora in these periods, it never reaches 0.5 per cent; per 1,000 words, it only reaches higher than 0.050 in the last subperiod. The percentages and numbers vary between the subperiods, making it difficult to determine whether there is an increase over time, except for one aspect: all periods after E1 show a higher percentage and number than E1, suggesting that at least it becomes more of an established construction. The low frequencies are not surprising: the recipient passive is even more restricted lexically than the prepositional passive and within the range of ditransitive verbs there are some that for a long time only allow a theme passive and no recipient passive. Moreover, the stylistic advice against the recipient passive may have slowed down the development and this influence may have continued to well in the twentieth century—the PPCMBE only has texts up to 1914.

As expected because of the limited number of ditransitive verbs, the range of verbs that occur with a recipient passive is quite small and remains largely similar over time. The total number of verbs found with recipient passives in the corpus is 27, distributed over 84 instances. Of these verbs, 13 only occur once while there are 5 verbs which occur in almost all subperiods: *allowed* (13 instances in total), *offered* (11 in total), *paid* (7 in total), *told* (7 in total), and *shown* (6 in total). Of these, *paid* and *shown* are among the earliest examples (both from E1), illustrated in (23) and (24).

- (23) *and he was shud evere plasse ther,*
 ‘and he was shown every place there’ (machyn-e1-p2,79.331)

- (24) *It was orsemen that al the light orsemen of Bolein and the men of armes shuld be paid their wages,*
(edward-e1-h,268.129)

While most verbs remain stable, this corpus contains one type of expansion of recipient passives to a new verb: even though there was an example of a recipient passive with *give* in the ME corpus (see Section 3.2.2.2), the first occurrence after that is in B3, followed by another 3 instances in the same subperiod.

While the recipient passive gradually becomes more established, the theme passive shows an opposite development, with decreasing frequencies, a decreasing range of verbs, and limited grammatical possibilities. The following table presents the frequency of the theme passive in the corpora.

Table 6.2 Frequency of theme passives in the PPCEME and PPCMBE.

Period	Theme passives	All passives	% Theme passives	No. of words	Theme passives per 1,000 words
E1	30	6,282	0.48	567,795	0.053
E2	54	6,545	0.83	628,463	0.086
E3	35	5,741	0.61	541,595	0.065
B1	21	3,528	0.60	298,764	0.070
B2	18	4,797	0.38	368,804	0.049
B3	3	3,330	0.09	281,327	0.011

Against all passives: E1-E2: $p < .05$; E2-E3: $p = .16$; E3-B1: $p = .93$; B1-B2: $p = .15$; B2-B3: $p < .05$

The table shows that in the first four subperiods the theme passive is more frequent than the recipient passive, with percentages around or above 0.5 per cent of all passives and above 0.050 instances per 1,000 words. In B2, the percentage of theme passives first drops well below 0.5%, but it is still more frequent in that period than the recipient passive, which has only 0.21 percent. Only in the last subperiod does the theme passive become really infrequent. Again, this shows that the status of the recipient passive as the standard may only have been reached quite recently, as a result of which it is not yet visible in the MBE corpus.

While the number of examples is much higher than for the recipient passive (161 vs 84), the range of verbs that occur with the theme passive is, not surprisingly, roughly similar: 32 verbs. There are some differences, however. While *give* hardly occurred with recipient passives in the corpus, it is the most common verb in theme passives with 38 examples, and in addition the only verb to occur in all subperiods. Other common verbs are *told* (22 over 4 periods), *offered* (13 over 5 subperiods), and *shown* (10 over 3). Interestingly, *told*, *offered* and *shown* were also among the most frequent verbs with the recipient passive, showing that for these verbs, there is real variation between the two alternative passives. With the exception of *give* and

offered, these verbs start to show gaps in the later subperiods, with the last occurrence of *told* and *show* occurring in B1. The remaining verbs in B3 are *give* (twice) and *owed* (which does not occur in earlier periods).

Another limitation for theme passives is that they overwhelmingly occur with pronominal recipients from the first period onwards, as previously discussed by Allen (1995: 432-5), who refers to Oehrle (1979) for this observation. As discussed in Section 3.2.1.1, pronominal ditransitives still show more possibilities for active and passive formation in PDE. Of the total of 161 theme passives in the PPCEME and PPCMBE, 148 examples (91.9%) contain a pronominal recipient, as illustrated in (25).

- (25) *it **was given me** by the soldiers after the Crimean War;*
(b3,nightingale-189X,437.396)

The last two occurrences of the total of 13 examples with nominal direct objects occur in B1, given in (26) and (27).

- (26) *for an assurance of both **was given the Prophet** at the same time.*
(b1,burton-1762,1,8.53)

- (27) *It were to be wished, that either a longer continuance **was allowed such men** at the hospital, or that their cure was rendered more perfect by a sweating course.*
(b1,lind-1753,243.32)

This fact could partly be explained by the general tendency for direct objects in active sentences to be nominal and for indirect objects to be frequently pronominal (since these are recipients, they are presumably often animate as well). Not taking into account the nominal or pronominal character of the other object, direct objects are only pronominal in 2.2% of the ditransitive clauses in the PPCEME and PPCMBE (81 of 3732 active ditransitives with two nominal objects). In contrast, indirect objects are pronominal in 85.2% of the ditransitive clauses (3181 of 3732).⁹⁴ Regardless of the explanation, the dominance of pronominal objects among theme passives signals a grammatical limitation of the theme passive, which must be a factor in the decreasing number of examples.

The discussion so far has already shown that the recipient passive remained very infrequent throughout most of the subperiods, while the theme passive seems to have been the more stable construction during this time. Figure 6.1 and Table 6.22 illustrate the balance between recipient and theme passives for each

⁹⁴ This was based on Query 6.2 for active clauses, with the stipulation that the IP* should dominate both a first object (direct object) and a second object (indirect object). A further query was then added to select (i) pronominal first objects and (ii) pronominal second objects.

subperiod. The lines in Figure 6.1 represent the percentage that one type of passive contributes to the total number of passives of ditransitives found in the corpora.

Figure 6.1 Distribution of theme and recipient passives in the PPCEME and PPCMBE

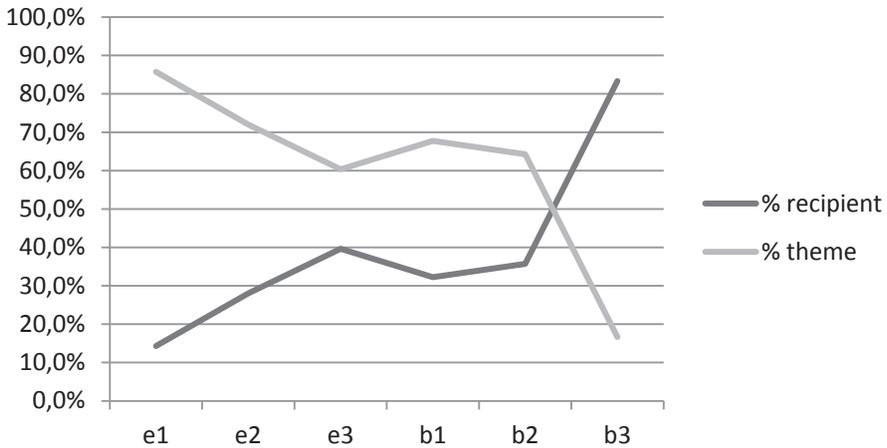


Table 6.22 Distribution of theme and recipient passives in the PPCEME and PPCMBE

		E1	E2	E3	B1	B2	B3
Recipient passives	<i>n</i>	5	21	23	10	10	15
	%	14.3	28.0	39.7	32.3	35.7	83.3
Theme passives	<i>n</i>	30	54	35	21	18	3
	%	85.7	72.0	60.3	67.7	64.3	16.7
Total	<i>n</i>	35	75	58	31	28	18
	%	100.0	100.0	100.0	100.0	100.0	100.0

Although the number of examples is quite small, these data clearly show that the theme passive remains the most frequent option of the two possibilities until the B2 period, after which it suddenly drops from 64.3% to only 16.7%. Two individual verbs are especially telling in this respect: *given* only occurs in B3 for the first time in a recipient passive, and *told* occurs as a theme passive from E1 to B1, and from E2 onwards in every single period with recipient passives. The shift in balance, then, only takes place in the final subperiod, further supporting the earlier findings that the recipient passive only slowly became a stable construction, and that the loss of the theme passive was also a gradual development.

In conclusion, the data in this section have not confirmed Visser's (1973) and Seoane's (1999b) earlier claims that the recipient passive showed a similar increase

over the EModE and Modern English as did the prepositional passive. While the recipient passive does become slightly more frequent after E1, there is no consistent increase and the percentages remain low. In contrast, the theme passive remains a much more frequent option throughout these periods. It is only in the B3 period that a change takes place: all verbs now seem to accept the recipient passive, and the recipient passive becomes slightly more frequent overall. The development of the recipient passive, then, does not reflect the general increase in use of passives over the course of the EModE and MBE periods and its true rise towards becoming the default option (in standard English) is likely to have taken place after the last subperiod of the MBE, in the twentieth century.

6.3.3 ECM passive

The introduction of the ECM presents a different scenario from the introduction of the prepositional and recipient passives because it is not just the passive that was new, but the entire ECM construction was introduced into the language. In addition, over the centuries the distribution of actives and passives has been described as ‘skewed’, with the passive being a more accepted and frequent ECM construction than the active. It was specifically the ECM passive that Los (2005, 2009) first proposed to be introduced as a consequence of the loss of V2, because of its ability to create unmarked subjects (see Mair 1990). The data in Chapter 3 showed that there were only a handful examples in the ME corpus and that there were passives that were structurally close—object control passives and fixed adjectival passives—which could have formed a model for the ECM passive, leading to acceptance specifically of the passive version of the ECM. This section traces the establishment of the ECM passive in the PPCEME and PPCMBE, investigating a proposed grammaticalization of the ECM passive and the distribution of active and passive.

One type of explanation that has been given for the skewed distribution of active and passive in PDE and earlier periods is Noël’s (2001) proposal for the grammaticalized status of the ECM passive in PDE. The grammaticalized status of the ECM passive—although described as a “weak kind of grammaticalization” by Noël (2001: 290)—means that it is separated in its development from the active, and that it has a specific function which may explain its dominance over the active version. According to Noël, the ECM passive has become more “schematic” over the centuries, allowing other verbs to be inserted in the pattern. Noël (2008) is also one of the few studies tracing the development of the ECM passive between the introduction and the present day. He provides data for the frequency of verbs used in the ECM passive in the period from 1640 to 1920 (CLMET corpus, *Corpus of Late Modern English Texts*, De Smet 2005), and from the *British National Corpus* (BNC). He finds that there is a decrease in the CLMET corpus, from 171.89 instances per million words in the first period, 1640-1710, to 107.73 in the period 1850-1920; however, the BNC data show a much higher frequency, with 221.60. Noël’s explanation for this is that the CLMET corpus contains mostly fictional texts (330, referring to De Smet 2005: 72), while the BNC presents a wider selection of genres,

including genres for which Noël in an earlier publication (2003) found that the ECM occurred more frequently, namely journalistic and science text (2008: 331). In other words, while the ECM may have become an established construction, it seems to be limited both in terms of the verbs with which it occurs and the genres in which it occurs.

The following sections present the results of a corpus study of the ECM active and passive in the PPCEME and PPCMBE. While we saw in the ME period that the related passives of object control structures and the fixed adjectival passives played a role in the introduction of the ECM passive, this section only focuses on the ECM constructions, because we already know that these passives continue to be used in PDE and once the ECM passive is introduced, their influence on the ECM passive is likely to be smaller.

Selection

At least a considerable portion of ECM passives should be selectable in the corpora according to the following criteria: a finite main or sub clause containing a form of *to be*, a passive participle, and an infinitival clause with an empty subject position, as in (28).

- (28) [_{IP-MAT} [_{NP-SBJ-1} *he*] *is said* [_{IP-INF} [_{NP-SBJ} *-1] *to be very clever, and very sensible.*]
(b3,victoria-186X,1,130.351)

The subject of the finite clause and the empty subject position in the infinitive are co-indexed, indicating that the subject of the infinitive has moved to the subject position of the main clause in this passive clause. These clauses were selected using the following query.

Query 6.8 ECM passives in the PPCEME and PPCMBE
a. IP-MAT* IP-SUB* idoms *SBJ*
b. IP-MAT* IP-SUB* idoms *BED* BEP* BEN BAG BE
c. IP-MAT* IP-SUB* idoms *VAN*
d. *BED* BEP* BEN BAG BE precedes *VAN*
e. IP-MAT* IP-SUB* idoms IP-INF*
f. *VAN* precedes IP-INF*
g. IP-INF* idoms NP-SBJ*
h. NP-SBJ* idoms **
i. IP-INF* idoms TO

The condition that the infinitival clause contains an empty subject position is necessary to separate ECM passives from object control (and other) structures, as illustrated in (29).

- (29) *And first, they began with Mr. Udall; [who, [_{IP-SUB} after he was called, was commanded [_{IP-INF} to stand aside till anon]].*
(judall-e2-p2,1,177.375)

In (29), the infinitival clause does not contain a subject node. However, this distinction is not fully implemented in the corpora, witness example (30), which is an ECM passive, but without a subject node in the infinitival clause.

- (30) [_{IP-MAT} *Lord Melbourne is reported [_{IP-INF} to have said some years earlier, when a similar opportunity had come to him, that he had no need to bribe himself]];*
(b3, trollope-1882,174.301)

In addition, the results of Query 6.8 included object control structures, which meant that all results needed to be analysed by hand. The relevant examples are discussed below, supplemented with examples of ECM passives which I came across in the corpus in different searches (for instance for the recipient passives).

Results and discussion

Table 6.23 presents the frequency of the ECM passive in the corpora, given as before both against the number of passives in each period and against the number of words.

Table 6.23 Frequency of ECM passives in the PPCEME and PPCMBE

Period	ECM passives	All passives	% ECM passives	No. of words	ECM passives per 1,000 words
E1	38	6282	0.60	567795	0.067
E2	91	6545	1.39	628463	0.145
E3	120	5741	2.09	541595	0.222
B1	35	3528	0.99	298764	0.117
B2	99	4797	2.06	368804	0.268
B3	50	3330	1.50	281327	0.178

Against all passives: E1-E2: $p < .001$; E2-E3: $p < .005$; E3-B1: $p < .001$; B1-B2: $p < .001$; B2-B3: $p = .07$

Judging from Table 6.23, the ECM passive becomes more established after the first subperiod, reaching 1.39% of all passives and 0.145 per 1,000 words in E2, after which it continues to show either similar or somewhat higher numbers, with the exception of B1. The table shows no clear development over time, with quite considerable variation between subperiods, ranging from 0.99% and 2.09% in terms of percentages and from 0.117 and 0.268 in terms of occurrences per 1,000 words. In both cases, there is a considerable decrease from E3 to B1, which may be caused

by differences between the PPCEME and the PPCMBE corpora: the PPCEME is a larger corpus and may present a slightly different selection of genres; the ECM in particular may be sensitive to this. Overall, the ECM passive does not seem to be a widespread construction, with examples in E1 found in only 13 of the 71 texts in the corpus. There is also still evidence of influence from Latin, with 13 examples of the 38 in E1 originating from a translation of *Boethius*. The restricted use of the ECM passive is also visible in the over-representation of certain texts in the corpus: for instance, in E3, two texts (*Boethius* and *Walton*) account for more than a quarter of all examples (33, 27.5%) and in B1, one text (*Priestly*) accounts for more than one third of all examples (12, 34.3%).

In addition to the somewhat small number of texts that contain ECM passives, the ECM passive is used with only a limited range of verbs. The number of times each verb occurs per subperiod varies considerably, ranging from 2.2 to 4.8, but it is more than 4 in three periods. This is considerably higher than for instance the prepositional passive, where this number ranged from 1.6 to 2.1, and can be seen as evidence for it being a fixed pattern or template, occurring only with a limited range of verbs. There are a number of verbs that occur in all six subperiods and account for many of the examples: *think*, *say*, *find* and *know*. Their frequency differs across the periods, but *say* always occurs as one of the three most frequent verbs—in E2 and E3 as the single most frequent verbs, with a considerable difference with the runner-up: 28 instances of *say* in E2, with only 12 for *know* in second place, and in E3, 24 instances of *say*, with *think* in second place with 13 instances. Interestingly, the prototypical ECM passive verb, *believe*, only occurs infrequently in this corpus, with 1 example in E3 and 1 in B3.

Most of the examples in the first periods seem to be what Noël defines as regular passives; in other words, these cannot be classed as belonging to the group with an evidential reading that develop into a fixed pattern (Noël 2008: 317-318, see Section 3.4.2.3). The most important characteristic for the regular passive is that it refers to an actual speech act, and that the person responsible for this act should be identifiable. In contrast, an evidential reading of ‘hear-say’ or of “shed[ding] responsibility for the truthfulness” (Noël 2008: 318). Consider the following examples.

- (31) *Attorney. In this Bill the Duke is proved to be 40,*
(thoward2-e2-p2,105.412)
- (32) *Truely, all these thinges haue ben shewed before, to be all one thinge.*
(boethco-e1-h,76.356)
- (33) *or if any persone were perceiued to be absent, or were sene to laughe at the folye of the emperour, he was forthe with accused, as it were, of missprision: wherby the emperour founde occasion to committee him to prison or to put hym to tortures.*(elyot-e1-h,27.56)

In example (31), the ‘speech act’ is given in the sentence itself, *this Bill*, while in example (32), the only possible interpretation is that there has actually been someone who has shown that *these things* are *one thing*, even though this person is not mentioned. In example (33), finally, there is the possibility of an evidential reading if only the first sentence is considered (*sene to laughe*), but in the following sentence it becomes clear that the author refers to a situation where there are actual people seeing these persons laugh: there is a consequence for these persons, so they must have been seen; it is not a hypothetical situation that is described, or a situation of hear-say.

The evidential reading that is characteristic for the ECM passive in PDE does not dominate the use of the ECM passive in the first subperiods, but increases in frequency in later subperiods. In the E1 period, there are no clear examples of evidential ECM passives. The first examples of a passive with an evidential reading are found in E2, given in (34) and (35).

(34) *The reputed Author therof **is said to bee** a famous practizer in Chirurgery, dwelling at Mountpelier in France*
(clowes-e2-h,31.282)

(35) *hee **is thought to bee** woorth twenty English millions:*
(coverte-e2-h,37.156)

In these two examples the author himself does not describe a specific ‘speech act’ or take responsibility for these observations. Rather, he puts the responsibility for the opinions with some unmentioned person. In later periods, there are more examples of evidential readings, as for instance the following examples.

(36) *If a don is crusty and silent he **is held to be** arrogant;*
(b3, benson-190X,113.261)

(37) *Thoughts of yellow fever and the land-crabs, which **were said to eat** the bodies of people who were buried there, occurred to me;*
(b3, fayrer-1900,9.223)

In the MBE period, the evidential reading accounts for more of the examples, although regular passives can also still be found—as indeed Noël (2001) has already shown in his definition of these types of ECM passives based on PDE data. Interestingly, the increase in examples with a clear evidential reading coincides with the later stages in the grammaticalization of the modals, from full verbs in OE to auxiliaries with limited syntax and morphology in EModE and MBE, and a designated function of marking modality (cf. Lightfoot 1979, Warner 1993).

Another observation with respect to the type of verbs that occur in the ECM passive is that the passive examples contain some verbs that are unexpected from a

PDE perspective. These were included because the selection was not based on specific verbs but on the structure of *to be*-past participle-infinitival *to*; the only examples that were excluded were clear object control structures. The fact that there are other verbs that occur in this pattern suggests that the ECM passive was not yet limited to the verbs of thinking and declaring. Consider the following examples.

- (38) *for certainly, a body that has so many pores in it as this **is discover'd to have**, from each of which no light is reflected, must necessarily look black,...*
(hooke-e3-p1,101.27)
- (39) *S. Oswaldes **is countid to be** auncient.*
(leland-e1-h,73.101)
- (40) *The Muskles of the neck, after Galen, **are numbred to be** .xx. mouing the head and the necke.*
(vicary-e1-p2,46.75)

The verbs in these sentences do not appear to have meanings of thinking and declaring (cf. the list from Postal (1974) and Los (2005) in (44) and (45) below) and are unlikely to occur with a *to*-infinitive in PDE. These examples, which represent regular passives in Noël's terms, show that the evidential use was not yet in place; the evidential reading mostly relies on the semantics of verbs of thinking and declaring.

Finally, let us consider the active ECM constructions in the two corpora. Technically speaking, there should be a difference in annotation between an ECM construction and an object control structure, as follows.

- (41) [_{IP-MAT} *The Author of this Book*, [_{VBP} *shews*] [_{IP-INF} [_{NP-SBJ} *this Method of Education*] *to be Antient and Infallible*]],
(b1,anon-1711,7.65)
- (42) *But let us now discourse a little upon the few things which [_{IP-SUB} [_{NP-SBJ} *our feeble reason*] [_{VBP} *permits*] [_{NP-OBJ} *us*] [_{IP-INF} *to know of the profound abyss of the Divinity*].]*
(b2,boethri-1785,173.417)

In (41), the subject of the infinitive is annotated as being part of the infinitival clause, while in (42), the subject of the infinitive is coded as being outside the infinitival clause, as an object of the finite clause. However, it seems that some ECM examples are also coded in this second way, witness example (43).

- (43) *For if their Masters perfectly understood the Latin Tongue, as [_{IP-SUB} [_{NP-SBJ} !] [_{VBP} suppose] [_{NP-OB1} them] [_{IP-INF} to do]] and always talk'd with them in pure Latin, they cou'd not fail of talking it Elegantly likewise.*
(b1,anon-1711,8.72)

For this reason, I used queries to select both types of sentences and added a requirement on the type of verb, selecting specific ECM verbs. Two groups of verbs were used. The first group of verbs is based on Los's list of verbs of thinking and declaring, which in turn is based on Postal (1974), repeated in (44). The second group is based on the verbs that were found with passive ECM, given in (45).

- (44) *Acknowledge, admit, affirm, allege, assume, believe, certify, concede, consider, declare, decree, deduce, demonstrate, determine, discern, disclose, establish, estimate, feel, figure, find, gather, grant, guarantee, guess, hold, imagine, intuit, judge, know, note, posit, presume, proclaim, prove, reckon, recognize, remember, report, reveal, rumour, say, show, specify, state, stipulate, suppose, surmise, take, think, understand, verify*
- (45) *Perceive, see, count, deem, conceive, adjudged, accounted, annal, construed, denied, discovered, esteemed, fabled, have, heard, mean, mention, number, observe, own, record, repute, suspect*

Queries 6.9 and 6.10 below were used in combination with this list of verbs.

Query 6.9 Active ECM in the PPCEME and PPCMBE

- a. IP-MAT*|IP-SUB* idoms *SBJ*
- b. IP-MAT*|IP-SUB* idoms *VBD*|*VBP*|VB
- c. IP-MAT*|IP-SUB* idoms IP-INF*
- d. IP-INF idoms NP-SBJ*
- e. IP-INF* idoms TO

Query 6.10 Active object control in the PPCEME and PPCMBE

- a. IP-MAT*|IP-SUB* idoms *SBJ*
- b. IP-MAT*|IP-SUB* idoms *VBD*|*VBP*|VB
- c. IP-MAT*|IP-SUB* idoms IP-INF*
- d. IP-MAT*|IP-SUB* idoms NP-OB*
- e. *VBD*|*VBP*|VB precedes NP-OB*
- f. NP-OB* precedes IP-INF*
- g. IP-INF* idoms TO

The difference between the two queries is that the ECM query selects clauses with a subject in the infinitival clause, while the object control query selects clauses with an

object in the finite clause. The results were analysed by hand, but I made no distinction between different uses of verbs, for instance *find* (actual finding versus opinion). Table 6.24 presents the frequency of the relevant examples.

Table 6.24 The frequency of passive and active ECM in the PPCEME and PPCMBE

Sub period	Passive ECM	Active ECM	% Passive ECM	No. of words	Passive ECM per 1,000 words	Active ECM per 1,000 words
E1	38	43	46.9	567795	0.067	0.076
E2	91	149	37.9	628463	0.145	0.237
E3	120	146	45.1	541595	0.222	0.270
B1	35	54	39.3	298764	0.117	0.181
B2	99	62	61.5	368804	0.268	0.168
B3	50	36	58.1	281327	0.178	0.128

For the balance active/passive across the subperiods: $\chi^2(5)$ 28.29, $p < .001$

Interestingly, this table shows that the passive and active are in equal distribution in the first four subperiods, with the passive representing between just over 40 and just over 50 per cent of the occurrences of these verbs. This distribution changes in the B2 and B3 period, where the passive makes up over 60 per cent of the instances. The last column shows how the active, after an increase in E2 up to 0.237, gradually becomes less frequently attested in the corpora, consistently decreasing to 0.128 in B3. These results present a refinement of earlier findings of an even more skewed distribution of active and passive, but the ECM nevertheless presents a markedly different distribution of active and passive than the overall corpora, with the average being roughly 20% passives. Moreover, individual verbs may behave differently: among the most common verbs in active and passive, some are used virtually only in the passive (for example, *say*: 1 active and 96 passive examples in the two corpora), while others occur more often in the active than in the passive (for example, *know*: 73 active and 30 passive).

In conclusion, the ECM passive gradually becomes a more stable construction, with some indications of initial restrictions of genre and Latin influence. The evidential use that is often found in PDE also seems to have developed over time, because it is not attested in the first subperiod. Interestingly, the data show that, contrary to expectations from a PDE perspective, there is still quite a high number of active ECM examples in the corpus, although the distribution of active and passive in the ECM passive is decidedly different from regular active/passive distribution in the corpora. Finally, the development towards an even greater dominance of the passive in the ECM construction, based on use with the verbs of thinking and declaring, only seems to take place in the Modern British period.

6.3.4 Summary and conclusions

This section has examined how the new passives—prepositional, recipient and ECM—became a fixed part of the language, and has considered the development of these passives against the general development of the passive as described in the previous section. The three passives all become more stable constructions, in the sense that they all show an increase in the early EModE period, and continue to be used at roughly the same rate from then on, with some showing an increase in the last period. All three passives are lexically restricted by the availability of a specific type of argument structure in the active, which means that there is a limit to the number of verbs with which they occur and to how frequent they can become.

After their contemporaneous joint introduction in the fourteenth century, the new passives mostly show individual stories of development. In the case of the prepositional passive, the increase in frequency is consistent over the periods, and is similar to the overall development of the passives. While there are some signs that the prepositional passive is grammatically restricted, there are more signs that it is a fully productive pattern which actually increases in possibilities over time. The recipient passive, in turn, only shows a minor development towards establishment in the first subperiods, and it remains a marginal construction until there are some signs of an increase in the final part of Modern British period. In that same period, it also, for the first time, is the most frequent passive of a ditransitive with two nominal complements, as the theme passive decreases in frequency. The ECM passive, finally, also shows an initial increase, although there is no consistent increase over the following centuries and percentages continue to vary considerably between subperiods (and texts). Interestingly, the active ECM remains quite frequent in the corpus, despite the general frequency of the passive version. In conclusion, the real changes in the use of these new passives seem to take place only in the last periods, which is slightly later than the general development of the passive. Nevertheless, it still falls into the expected periods for the responses to the developments in the clause-initial position and the need for subjects.

Section 6.4 Middles and non-agentive subjects

This section broadens the perspective and investigates two strategies—middles and non-agentive subjects—which, like the passive, function as subject creators, and can be seen as another response to the reliance on subjects for encoding unmarked themes after the loss of V2 and hence to a need for subject strategies. These strategies make English stand out both with respect to its sister languages, and from a cross-linguistic perspective. Previous researchers have ascribed the occurrence of (some of) these strategies to the rigid SVO order in English (Kirkwood 1978), to the loss of case (Hawkins 1986), and more recently to the loss of V2 (Los 2009, Los & Dreschler 2012). An important question for the current study is not just whether these strategies function as subject creators in PDE, but also whether these

strategies are innovations in the history of English. In this section, I present evidence from the literature and the annotated corpora for the claim that they are indeed post-V2 innovations.

The most frequent type of middle (see Keyser & Roeper 1984, Hundt 2007), as illustrated with some recent examples in (46)-(47), uses an active intransitive verb in combination with an adverb, resulting in a passive meaning.

- (46) A: It seems to be impressing our American friend.
B: Americans impress easily
(Inspector Lewis, Series 2, Episode *And the Moonbeams Kiss the Sea*)
- (47) I think [Kate Middleton] photographs better than she knits.
(Graham Norton Show, Series 9, Episode 1, 15 April 2011)

The meaning of these sentences is passive: Americans *are easily impressed*; she *is photographed better* than she *is knitted*, i.e. she looks better on a photograph than as a knitted figure. What makes the middles relevant for the present investigation is that they create non-agentive, and hence non-canonical, subjects and as such represent an additional strategy in English to create subjects. In addition to the middles, it has been observed that PDE subjects are more generally less restricted to the semantic role of agent. Rohdenburg (1974) presents an extensive collection of English examples which contain a non-agentive and inanimate subject, as in (48) and (49), which are impossible in German.

- (48) This hotel forbids dogs.
(49) This tent sleeps four.
(Rohdenburg 1974, as quoted by Hawkins 1986: 58, his 4.10 and 4.17)

Several authors, including Rohdenburg (1974), Hawkins (1986) and Toyota (2008), have claimed that both middles and the use of these non-agentive subjects are innovations in the history of English, although the evidence for this claim is limited.

Previously, authors have explained the existence of these constructions with reference to the strict SVO order in English, but more recently Los (2009) and Los & Dreschler (2012) have connected the development to the loss of V2. Kirkwood (1978: 239), for instance, claims that the non-agentive subjects, like passives, are a response to the strict SVO word order in English, as a means to reorder information in the sentence to achieve given-before-new (what he calls theme-rheme). Hawkins (1986) also takes up this idea, adding that the loss of case is ultimately the cause, because it led to the strict SVO order. However, these authors overlook the position of Dutch—a closely-related language, but one which has lost the major case markings, like English, and unlike English, has not acquired a strict SVO word order. Los (2009) and Los & Dreschler (2012) include non-agentive subjects in the range of consequences of the loss of V2, building on the need for subject strategies.

The section is structured as follows. Section 6.4.1 aims to define the prototypical middle and some related constructions, and attempts to draw a diachronic picture for the occurrence of middles. Section 6.4.2 discusses in more detail the non-agentive subjects and again aims to find evidence for the claim they are innovations in the history of English. Section 6.4.3 concludes.

6.4.1 Middles

This section provides a definition for a range of middle constructions and presents a syntactic analysis that relates them structurally to passives and unaccusatives. I then attempt to substantiate the proposal put forward in the literature (e.g. Visser 1973, Kirkwood 1978 and Hawkins 1986) that these middles are ME or EModE innovations, in order to establish to what extent the date of their introduction and development provides evidence for the hypothesis that they are another response to the need for subjects after the loss of V2.

6.4.1.1 Defining middles

The key characteristic of middles is that they have active morphology but a passive interpretation, as illustrated in (50) with a what I will call ‘prototypical’ middle, which has an intransitively used verb in combination with an adverb.

(50) Bureaucrats bribe easily.

(Keyser & Roeper 1984:383)

In addition to the type of middle in (50), there are other types, for instance without an obligatory adverb, and consequently there is some debate on the defining characteristics of middles (see Ahn & Sailor to appear and Hundt 2007 for an overview). There nevertheless seems to be consensus in the literature on at least two aspects of middles: the non-agentive nature of the subject and the relation to verb transitivity alternations.

Ahn & Sailor present the non-agentive nature of the subject as one of the “exceptionless” properties of middles, based on the literature on middles since Keyser & Roeper’s (1984) early publication. This property relates middles to passives, as illustrated in the following examples.

- | | |
|--|-----------|
| (51) a. Mobsters bribe bureaucrats easily. | (Active) |
| b. Bureaucrats _i bribe <i>e.c._i</i> easily. | (Middle) |
| c. Bureaucrats _i are bribed <i>e.c._i</i> easily. | (Passive) |

(Ahn & Sailor to appear: 5, their 17)

The internal argument in the active, *bureaucrats*, occurs as the grammatical subject both in the middle in (51)b and in the passive in (51)c, leaving a trace in its original

position—shown by Ahn & Sailor’s “e.c.” for empty category. This also means that middles cannot be used with agent-oriented adverbs, as illustrated in (52).

- (52) a. Students always translate Greek begrudgingly. (Active)
 b. Greek is always translated begrudgingly. (Passive)
 c. *Greek always translates begrudgingly. (Middle)
 (Ahn & Sailor to appear: 6, their 22)

The second property of middles is their similarity to transitivity alternations: like the standard intransitive/transitive distinction—as illustrated in (53)—middles take a verb which is traditionally a transitive verb, and use it intransitively.

- (53) a. The sun melted the ice.
 b. The ice melted.
 (Keyser & Roeper 1984: 381, their 2)

For Keyser & Roeper, following Burzio (1981), ergatives such as (b) are lexically intransitive (similar to intransitives like *sleep*), i.e. they do not have an internal argument when they are base-generated. Keyser & Roeper propose that middles, in contrast, are lexically transitive: they are generated initially with an internal argument, but this internal argument is deleted in the syntactic derivation. Hundt (2007) presents additional examples where the middle takes an typical transitive verb like *shock* in (54)a, but uses it intransitively, as in (54)b.

- (54) a. It did not shock me.
 b. After almost four years in the Senate, Barbara Boxer doesn’t shock easily.
 (*Newsweek*, 8 September 1997, p. 38)
 (Hundt 2007: 8, her 3c)

In short, the middle allows verbs to occur intransitively which would not normally occur in an intransitive version.

In addition to these properties of non-agentive subjects and transitivity alternations, several other properties have been discussed, some of which also relate to the types of verbs used in middles. For instance, Ahn & Sailor translate the condition on the verbs by saying that the verbs used in middles must be causative, as shown in (55), and non-stative, as shown in (56).

- (55) a. Causative: You can easily {raise/*rise} this flag if you use a pulley.
 b. Middle: This flag {raises/*rises} easily if you use a pulley.
 (56) a. Active: You can easily {store/keep} all the tools in the shed.
 b. Middle: All the tools {store/*keep} easily in the shed.
 (Ahn & Sailor to appear: 6, their 24a-b and 25a-b)

The middle in (55)b is only grammatical with the causative version of the verb, *raise*, while in (56)b, it is only possible with the non-stative *store*. This may not be as strong a condition as presented by Ahn & Sailor, and may be problematic as a generalization beyond specific examples; other elements in this sentence may force a middle reading, such as the presence of *easily* in (55)b. Second, Hundt (2007) points to another construction to which middles are possibly related, namely reflexive constructions, as witnessed by the following examples. The middle of (57)a has an alternative expression in the reflexive (57)b.

- (57) a. Heat-sensitive material molds to your body shape.
 b. Uses body heat to mold itself to fit your shape exactly (*Sky Mall Catalog*, Spring 1998, p.37)

(Hundt 2007: 11, her 11a and 11b)

What makes the definition of middles difficult is that there are counterexamples to most formal properties. Hundt (2007) illustrates the versatility of middles—which turns out to be much greater than the discussions by Keyser & Roeper and Ahn & Sailor suggest—with examples that show that middle is not restricted to one particular range of verbs or grammatical contexts, as those in (58)-(61).

- (58) [...] medium-bodied with a long bracing, acidic finish. [...] Should *cellar* well for another three years ('Wine commentary, *Christchurch Star*, 1/5/1996; attested example from Yoshimura 1998: 155).

- (59) [...] sleepwear that *machine washes* (Sears & Roebuck, 1986: 272)

- (60) [...] the routes are designed to *bicycle* in a few hours (Frown, F41, 5f.)

- (61) These 85s *are drinking* well. (This Life, Episode The Plumber Always Rings Twice, BBC TV, 1997)

(Hundt 2007: 1-2, her 1e, 2d, 3a and 4a)

Example (58) shows that the middle is used with a range of verbs, also unlikely ones like *cellar* that cannot normally be used intransitively and are rare as verbs in all contexts. Example (59) shows that it can be used with compound verbs, while (60) shows that it can also be applied to verbs that are already intransitive. Finally, (61) shows that the middle can occur in the progressive, although Hundt notes that these are especially rare. These examples show both the productivity of the pattern and its versatility.

The properties of middles pose two questions for the syntactic analysis of these sentences: merge or non-merge of the external argument and presence or absence of a VoiceP (for a more complete overview, see Hundt 2007). First, as for the passive,

there is the question whether the external argument is originally merged or not. Keyser & Roeper (1984) argue that an external argument is originally generated, but that a clitic “absorbs” case and the agent θ -role, somewhat similar to Baker et al.’s (1989) analysis for the passive. Ackema & Schoorlemmer (1995) in fact propose that the grammatical subject is the external argument, so no deletion of arguments or movement of the argument takes place. As Ahn & Sailor point out, many of these accounts (also Stroik 1992, Fujita 1994 and Hoekstra & Roberts 1993) rely on lexical rules to explain the deletion of the external argument, and that way violate minimality conditions: if the external argument is merged, on top of the VP, following general principles, then the internal argument, merged inside the VP, needs to move past the external argument in order to move to the surface subject position (although Fujita 1994 provides a solution for this problem). As an alternative, Ahn & Sailor propose that the external argument is not originally merged in middles, and the internal argument moves to the subject position. This analysis is very similar to the analysis proposed for the passive and unaccusatives in Section 2.3.3 and I will extend this idea to the analysis of the middles.

A second discussion, also reminiscent of the discussion of the passive, is the proposal that a VoiceP is involved in the derivation of middles. Ahn & Sailor’s analysis makes use of a VoiceP on top of the vP, following Collins (2005) and Kratzer (1996); in Collins (2005), Spec,VoiceP is the base-generated position for all external arguments. Ahn & Sailor’s VoiceP is present in all clauses, but has three different possible ‘settings’: active, middle or passive. In an active sentence, what is merged in Spec,VoiceP is the external argument. In a middle construction, the entire vP moves to Spec,VoiceP, in the absence of an external argument. Recall from Chapter 2 that for the current work, a VoiceP was dismissed, in favour of a flavours-of-v approach.

I will take the basic analysis of middles to be similar to that of unaccusatives: an internal argument is moved to Spec,TP because the subject position needs to be filled; in the absence of an external argument, the internal argument is the closest available argument. This still leaves the question of what the difference is between the ergative in (62) and the middle in (63).

(62) The ice melted.

(63) Bureaucrats bribe easily.

The main difference in interpretation seems to be the passive reading (for lack of a better term) of the middle. I would like to argue that this difference is not syntactic. The syntactic difference lies solely with the obligatory presence of the adverb in specifically this prototypical middle. In both the active and the middle, the subject corresponds to the direct object of a transitive verb, *melting* and *bribing*. One source of difference between the two sentences can then lie with the semantic roles, of both the internal and external argument. Where the *bureaucrats* can arguably be said to be the PATIENT, the semantic role of *the ice* can perhaps better be described

as THEME. This requires a full inventory of roles beyond the number of examples discussed here.

6.4.1.2 New middles?

In addition to the prototypical middle, several other types of clauses have been described as middles. First, Ahn & Sailor propose that the following two constructions—which they call the *accommodation* (64) and the *make* construction (65)—should also be analysed as middles.

- (64) a. My car seats four people.
 b. This bed sleeps five people.
- (65) a. Clowns make good fathers.
 b. This photo makes a fun dartboard.

(Ahn & Sailor: 1, their 2a-b)

Note that these constructions contain verbs that can also be used transitively, like many of the prototypical middles. Ahn & Sailor's claim that these constructions are middles is based on the fact that they meet the requirements which they consider to be the "exceptionless" properties of middles, illustrated in the following examples.

- (66) a. *Two people are slept by queen size beds.
 b. *A lovely drag queen is made by him.

(Ahn & Sailor: 3, their 10a and 13a)

- (67) a. These extra-long beds sleep tall guests (*by hoteliers).
 b. Thomas will make a great janitor (*by his OCD).

(Ahn & Sailor: 7, their 28b and 29b)

Examples (66)a-b show that the constructions do not allow passive formation; examples (67)a-b show that an external argument cannot be expressed. However, while the ungrammaticality of these examples is clear, note that the impossibility of passivization does not seem to hold for all the verbs mentioned by Ahn & Sailor, witness example (68).

- (68) How many people can be seated in the 2012 London Olympic Stadium?
 (<http://wiki.answers.com/>)

Sleep in (66) above does not allow a passive because it only occurs (normally) as an intransitive (ergative). One of their other claims is also problematic, namely that "the surface subjects [in the *make* and *accommodation* constructions] bear theta-roles *inconsistent with external arguments*" (7, italics mine). They claim that external

arguments, i.e. subjects in transitive clauses, are normally agent, causer or experiencer, while these subjects are themes or locatives. It is true that the subjects in the *make* and *accommodation* constructions do not bear prototypical subject theta-roles, but the existence of a whole range of non-agentive subjects in PDE, such as for instance *This hotel forbids dogs*, shows that “inconsistent” is too strong a claim.

I would like to propose that both constructions do not in fact represent middles, but fall within the range of non-agentive subjects allowed in PDE. The *make* construction should be analysed as a type of copula construction with a non-agentive subject, while the *accommodation* construction is a regular transitive clause but also with a non-agentive subject. If they are analysed in this way, the other properties follow from that, in particular the claim that they cannot express another external/agent argument—because they already have an external argument—and the impossibility of passivization.

A second type of sentence that has been analysed as another additional type of middle is presented by Toyota (2008) for the “necessative passive”, as illustrated in (69).

(69) The TV needs fixing.

(Toyota 2008: 207)

Toyota describes this construction in terms of modality, which he claims to have “an important role in its development and existence in PDE” (2008: 211). Toyota does not use any tests to determine the syntax of the construction. One condition described in this section that applies straightforwardly is the impossibility of adding an agent-oriented adverb.

(70) *The TV needs fixing begrudgingly.

What is less obvious is the impossibility of adding an agent *by*-phrase. Although (71) below is clearly problematic, it is possible to find examples with a *by*-phrase, witness (72).

(71) ?The TV needs fixing by John.

(72) Politics is a job that needs doing—by anyone who is interested enough to train for it and work at it. It’s like housekeeping; someone has to do it. Whether the job is done by men or women is not important—only whether the job is done well or badly.

(US Representative Helen Gahagan Douglas, quoted in Wasniewski 2006: 234)

There is a break after “needs doing” in example (72), setting it apart from the rest of the sentence, but the *by*-phrase clearly represents an agent belonging to the verbs *needs doing*. And while the whole construction, or the main clause, cannot be

passivized, witness (73), it is possible to replace the *fixing* with a passive infinitive, see (74).

(73) *Fixing is needed by the TV.

(74) The TV needs to be fixed.

The test that the surface subject is an underlying internal argument is problematic because there are two clauses involved in *The TV needs fixing*, as the passive alternative *The TV needs to be fixed* shows. The subject is the internal argument of the lower clause '*fixing/to be fixed*', but it is not the internal argument of *need*. Interestingly, another type of construction exists where the lower clause is not a gerund but a participle as complement, as in (75)-(77), examples from the OED.

(75) *This lock needs sortit. The hail house needs guttit.* (1923)

(76) *Does my hair need combed?* (1954)

(77) *He hoped to goodness that Mother and Walter got back home before Benny woke up from his nap and needed fed.* (1992)

I would like to propose that these *needs fixing* examples can only marginally be considered as middles. The middle properties do not straightforwardly apply, and it seems to be that the middle properties are located in the subclause rather than the main clause: it is the subclause that lacks an external argument. The only available argument for the subject position of the main clause is the internal argument of the subclause, *the TV*. The different constructions, i.e. the alternation with the passive infinitive in the lower clause, are accounted for by differences in the morphology of the lower construction: the form of the non-finite verb, passive morphology or leaving out the auxiliary.

In conclusion, I have argued against analysing the *make* and *accommodation* constructions as middles, as proposed by Ahn & Sailor. While there are some questions about the status of the *needs fixing* sentences as middles, as proposed by Toyota (2008), I will include those in the corpus study because they at least contain some properties of middles and, in addition, represent a construction that can be easily selected.

6.4.1.3 Middles diachronically

Previous authors have claimed that the middle is an innovation, with many basing their claim on Visser's (1963-1973) data (such as Lightfoot 1979, Toyota 2008, Hundt 2007), but the actual data that are available to support this claim are limited. One of the reasons for the lack of data seems to be the fact that the middle has few unique syntactic features, which means that finding middles in a corpus is not

straightforward. However, work on related constructions (Van Gelderen 2011) as well as work based on different corpora (Hundt 2007 for Modern English, De Smet 2013a for Early and Late Modern English) shows that there is enough suggestive evidence to support the claim that the middle, if not an actual innovation, in any case increased in productivity from roughly the seventeenth century onwards, during the transition from the EModE to the MBE period, and continued to grow into the twentieth century. This section discusses the data studies presented in the literature for the prototypical middle and the *needs fixing* construction, supplemented where possible with data from the PPCEME and PPCMBE.

The canonical middle: *The book sells well*

The canonical middle, of the type *the book sells well*, is described in several early grammars, such as Sweet (1900: 90, §249), who calls them ‘passivals’, and Jespersen (1909-1949: §16.8), who presents a list of examples from a range of texts and periods, although not ordered diachronically. Visser (1963-1973) presents a more extensive list of examples. He describes the actions in these middles as “quasi-automatic”, as if there is no agent present (1963-1973: §163). The earliest example in his list is the following fifteenth-century sentence.

- (78) *Grete pleynte ... of Wynes made nygh the seide Portz come into this londe ... ate that tyme ... the tonne of such Wynes **solde better chepe by a gretter quantite** than it is now’ [MMED]. 1437, Rule Parlt. 5, 113b*

This first verb to occur in the list, *sell*, is also the most frequent in the list, with examples from the seventeenth, eighteenth, and twentieth century. Other early, i.e. pre-1700, verbs include *soil* (1530), *pull* (1641), *steer* (1669), *tell* (1584), *peel off* (1634) and *vend* (1622). The examples become more frequent from the eighteenth century onwards, although the total number of examples given by Visser is only 42. Interestingly, Visser includes *open*, one of the typical alternation verbs, which means that it can be used as an intransitive in sentences which are not middles (*The door opened*). Yet the example from Visser, given in (79), clearly has a middle reading, rather than a simple unergative reading, probably because of the adverb.

- (79) *It [sc. the book] **opened very easily** at a certain page.*
A. Christie, *Evil under Sun* (Pock. Bks) 163.

The OED gives additional types of examples with *sell*⁹⁵, but only from the seventeenth century onwards, as shown in (80)-(83). Note, incidentally, that *sell* is one of the few verbs for which this use is given separately in the OED.

⁹⁵ Visser also used the OED but did not include these examples.

- (80) 1656 There is no Merchandize in this Ware-House which sels better, then certain Fans.
- (81) 1833 They sell at about a shilling a dozen.
- (82) 1851 I found a bookseller to publish my treatise. It sold well.
- (83) 1896 This, madame, is selling very well.

Not all sentences above have the format of the prototypical middle with an adverb, but instead have different complementation patterns, such as *sell at a about a shilling a dozen* in (81).

These examples suggest that middles are already used during the periods covered by the PPCEME and PPCMBE. However, attempts to find these middles in the corpora proved unsuccessful. It is possible to select clauses which have a finite verb and an adverb, without an object, complement *that*-clause and passive participle, but this results in around 3,000 clauses in the PPCMBE which include many irrelevant examples. If only lexical verbs (and not auxiliaries and modals) are selected, and the adverb should immediately follow the lexical verb and should not be a temporal adverb (which are labelled separately), the result is around 200 to 250 examples for each subperiod, but again, these are mostly irrelevant examples, such as (84)-(85).

- (84) *The Master talked suavely, interestingly, continuously.*
(b3, benson-190X,107.75)
- (85) *the women squeaked feebly,*
(b3, benson-190X,111.173)

The verbs in these examples are inherently intransitive and the adverb only modifies the active action described by the verb, without changing the sentence into a middle. The final strategy that I employed was to combine a general search for sentences with a finite verb and adverb (and without an object or *that*-clause) with the verbs that Visser (§168) found as occurring with middles.⁹⁶ This query resulted in a small number of examples which may qualify as middles. Consider (86) and (87), among about 100 examples per subperiod which did not represent middles.

⁹⁶ The verbs are *soil, peel, tell, pull, steer, vend, polish, read, spoil, tear, thresh, wear, compose, fuse, let, load, make, milk, paint, photograph, plough, sing, smoke, subscribe, transcribe, transplant, wrap, write, open, scare, shock, tire*. I changed some of the forms to make sure all finite forms and past forms were included.

- (86) *When old, they sell to very good Account for an Inland Market.*
(b1, drummond-1718,37.415)
- (87) *Such cooks as excelled in the art of preparing dishes sold for four talents;*
(b3, long-1866,2,96.494)

In conclusion, while there is evidence from Jespersen (1909-1949), Visser (1963-1973) and the OED that middles already occur as early as the fifteenth and sixteenth century, there is little to no evidence of middles in the MBE corpus. Of course, the lack of examples in the PPCMBE does not mean that middles are not used during this time; it only means that they do not occur in the corpus, or that the method of selection was not sufficient. As a result, it is, at this point, not possible to provide more substantial evidence for the claim that the middles were EModE innovations or to build a picture of the development of these middles throughout the centuries.

Hundt (2007): increase in early twentieth century

One solution for the problems with selecting middles automatically in a corpus is to manually analyse samples of text, which is what Hundt (2007) has done. Her data are based on twentieth-century material and they provide an interesting insight into the behaviour of the middle, or mediopassive. Most importantly, she finds an increase in the use of the middle over the course of the twentieth century.

Hundt's study consists of a manual analysis of mail-order catalogues from the American company *Sears & Roebuck* from four different years: 1897, 1927, 1957 and 1986, following earlier observations that middles are especially frequent in advertising language. Table 6.25 shows the frequency of the middles in the corpus, presented together with a range of constructions which Hundt shows to be connected to the middle or relevant for its development. The selection for the related constructions is based on the verbs that are used in mediopassive constructions.

Table 6.25 Mediopassive and alternative constructions in the *Sears & Roebuck* catalogues (based on the verbs that are used in a mediopassive construction), (Hundt 2007: 160, her Table 6.3).

Year	Active	Passive	Medio-passive	Pass.inf.	Reflexive	Total
1897	578	584	431	150	8	1751
1927	248	655	499	167	7	1576
1957	2448	730	2670	690	15	6553
1986	1307	290	1119	67	6	2789
Total	4581	2259	4719	1074	36	12669

A comparison of the numbers for each construction in this table shows that the mediopassive is a frequently used construction, with 4,719 occurrences in total—somewhat higher than the active (4,581) and considerably higher than the passive (2,259). Hundt adds that she found 425 different verbs which make up the 4,719 mediopassive examples, which she takes as evidence for the productivity of the construction. Hundt provides an overview of the differences in the verbs that are used in the construction over time, with some going out of use (such as *make up*, see (88), 2007: 161), while others only start to appear in the later catalogues. One example of this is the new verb *zip*, illustrated in (89).

(88) Fabric that “makes up pretty in tailor-made skirts” (*Sears & Roebuck*, 1897: 305)

(89) Curduroy Slacks .. zip to fit each side of waist (*Sears & Roebuck*, 1957: 87)
(Hundt 2007: 162, her 37)

Calculated against the total number of examples with these verbs, Hundt finds that the mediopassive increases in the twentieth century, from 24.6% to around 40%, while the passive uses of these verbs decrease (2007: 162-3). She provides additional percentages based on the exclusion of four verbs, *sell*, *retail*, *clean* and *wash*, which, she observes, have a skewed distribution. This second set of percentages shows an even more convincing increase of the mediopassive, from 26.1% to 56.6%. Although the mediopassive outnumbers the occurrences in active transitive sentences in these periods, the selected verbs continue to be used frequently in active transitive sentences. This means that the middle is not based on a specific and limited range of verbs but that the middles are really part of an alternation system which allows for both active, passive and middle use of particular verbs.

In conclusion, Hundt’s data show that—at least in her selected genre—there is an increase in the use of the mediopassive construction in the twentieth century. While her study does not provide information about the development in the preceding centuries or the first occurrences of the middle, this increase does provide us with evidence of an increase in what seems to be the most relevant period in all studies presented in this chapter so far, the late Modern period to the PDE period. Hundt’s data, then, add to the general impression in the literature that the middle becomes a more frequent pattern.

The necessative passive: *it needs doing*

Let us now turn to the other type of middle identified in section 6.4.1.2, the type *it needs doing*. Toyota (2008) bases his overview of the diachrony of the construction on Visser’s data (1963-73: §1788). He identifies three groups of verbs in Visser’s data: (i) a group of verbs that are no longer used after the seventeenth and eighteenth century, such as *abide*, *avoid*, *continue*, *escape*, *lack* and *suffer*; (ii) verbs

that are only in the nineteenth and twentieth century, such as *await, miss, prevent, repay, stand*; and (iii) verbs that occur in the construction consistently from 1400 onwards, such as *bear, deserve, merit, need, require, want* (2007: 207-209). In his own study—based on parts of the Helsinki corpus, the ARCHER corpus, the London-Lund corpus and the LOB—he finds only 14 examples, all occurring in PDE.

De Smet (2013a) also discusses the *it needs doing* construction, which he does not describe as a middle, but as ‘passival gerund construction’. He relates it to another construction with a passive interpretation of the gerund, which he calls the ‘passival participle construction’, illustrated in (90). The difference with *the TV needs fixing* is that this construction has an additional direct object.

(90) She needs everything rearranging. (BNC)

(De Smet 2013a: 84, his 10d)

De Smet searches for these two constructions in several historical corpora: the *Old Bailey Corpus* (OBC), the *Corpus of Late Modern English Texts* (CLMETEV), the *Corpus of EModE texts* (CEMET), the *Oxford English Dictionary* (OED) and the collected citations in Visser (1963-73). He finds that the passival gerund construction occurs with *need* from the late fourteenth century and with *want* from the late sixteenth century onwards, with the first examples given in (91) and (92).

(91) (a1398) *Everich ere þe spraiē [of a vine] nedep kuttinge and paringe.*

(92) (1574) *It was but rough hewen by one of the prentises, and wanted sum polishing by the forman.*

(De Smet 2013a: 85, his 17a-b)

De Smet (2013a) shows that the *needs doing* examples were more productive in Early and Late Modern English than they are in PDE, showing a more elaborate range of verbs: *decline, escape, fear* and *prevent* (referring to De Smet 2013b: 146-147), in addition to Visser’s range of verbs. Part of his explanation for the earlier frequency of the passival gerund is that the gerund could still more easily lead to a passive as well as an active interpretation, while he ascribes the decline in later periods to an increase in of active, clausal gerunds (2013a: 86).

Let us turn to the PPCEME and PPCMBE, corpora not included in De Smet’s data and only partly in Toyota’s study. The *needs doing* construction has a more distinctive syntactic character than the canonical middle and can for that reason be straightforwardly selected in the annotated corpora. The following query was used.

Query 6.11 'It needs doing' in the PPCEME and PPCMBE

- a. IP* idoms *VBP*|*VBD*
- b. IP* idoms *OB*
- c. *VBP*|*VBD* precedes *OB*
- d. *OB* idoms N
- e. N doms *ing

Lines a and b select clauses with a finite verb and an object, while line c stipulates that the finite verb should precede the object. The gerunds are coded as nominal objects, so many of the results consisted of regular nouns ending in –ing (like *king* or *thing*), which meant that the results needed sorting by hand.

Table 6.26 presents the number of examples per period.

Table 6.26 'It needs doing'-examples in the PPCEME and PPCMBE

Period	E1	E2	E3	B1	B2	B3	Total
No. of hits	7	3	2	2	3	6	23

The table shows that the construction is not very frequent, with only a handful of occurrences in each subperiod, but, as expected, it is consistently present in all periods. There is a clear difference in the types of verb that occur in the EModE period and the MBE, with an increasing dominance of *want* and *need* in the later periods. In the EModE period, *want* does not occur and *need* only occurs 3 times, of the total of 12 examples. In total, there are 7 different verbs, with *lack*, as illustrated in (93), occurring most often, 4 times.

(93) *For as the saying is, it lacketh wethering*
(latimer-e1-h,24P.96)

(94) *Your letter escaped opening,*
(phenry-e3-p2,343.42)

In the MBE corpus, in contrast, *want* and *need* make up 9 of the 11 examples (7 and 2, respectively), with *bear* and *require* accounting for the two remaining examples.

In conclusion, the data from the PPCEME and PPCMBE confirm the earlier findings from Visser (1963-1973) and De Smet (2013a) that the *it needs doing* pattern already occurs from the early sixteenth century onwards and continues to occur throughout all periods. It also confirms that the pattern seems to have allowed more verbs in earlier periods but has moved towards a restriction in the MBE period, in that it now mostly occurs with *need* and *want*. In contrast to the prototypical middle, then, it is possible to provide a more specific timing for the introduction of the *it needs doing* pattern: they seem to be a late ME or EModE innovation, with the first examples occurring around 1400, the same time that the new passives were

introduced. In contrast to these new passives, however, it becomes a more restricted pattern over time, rather than growing in possibilities.

A note on valency in the history of English

The final piece of evidence for the proposal that middles are an innovation comes from Van Gelderen (2011), who investigates basic valency in West Germanic languages and takes the history of English as a case study. She does not discuss middles specifically, but investigates the related point of transitivity alternations. Van Gelderen shows that in the history of English there is an increase in what she calls, following authors such as Haspelmath (1993) and Comrie (2006), labile verbs: verbs that can be used both transitively and intransitively. Van Gelderen concludes that OE already had quite a considerable number of labile verbs compared to its sister languages, but that the number of exclusively intransitive verbs decreases: 223 OE intransitive verbs from Visser (1963: 100) become 30 or 40 in Modern English (Van Gelderen 2011: 119). Example (95) and (96) illustrate two intransitive verbs in OE, the first unergative and the second unaccusative.

- (95) *Ac ðonne we slapað fæste ðonne we ...*
 but then we sleep fast then we
 'But then we are fast asleep...' (*Pastoral Care*, Hatton 195.5; Visser 1963: 97)

- (96) *ða unrotnesse ðe ðæræfter cumað*
 the sadness which thereafter comes
 'the sadness which comes thereafter' (*Pastoral Care*, Hatton 187.15)
 (2011: 113, her 10 and 12)

In contrast, the number of labile verbs increases: Van Gelderen (2011: 122) shows that 80 OE labile verbs (based on Visser 1963: 98 and Hermodsson 1952: 196-207) become about 800 labile in Modern English, based on McMillion (2006: 223-232). Examples (97) and (98) illustrate an intransitive and transitive use of the OE verb *baðian*.

- (97) *Ond seldom in hatum baðum heo baðian wolde*
 and seldom in hot bath she bathe would
 'and she would seldom bathe in hot water'
 (Bede 4.19; Visser 1963:99, Miller's edition 1891:318.15-16)

- (98) *and wolde seld-hwænne hire lice baðian butan to heahtidum*
 and would seldom her body bathe except at high-holidays
 'and would seldom bathe her body except at high-holidays'
 (Aelfric, *Lives*; from Visser 1963:99, as in Skeat 1881:434.44-45)
 (Van Gelderen 2011: 120, her 20a-b)

Van Gelderen’s study does not address the periods between OE and PDE, and due to the nature of the question (large-scale) and the method (assembled observations from the literature), there is no possibility to pinpoint the change in a particular period. What is needed for such detail is a study of the development of individual verbs as well as a survey of verbs in ME and EModE. However, it is unlikely that this type of change—i.e. involving a high number of verbs in this way—takes place within the span of a couple of generations, making it unlikely that it was already in full swing in ME. Most importantly, what Van Gelderen’s overview shows is that there is a general tendency of an increase in labile verbs throughout the history of English, and it might very well be this development that the middles fit in—making use of this tendency and further increasing the tendency by opening a whole new range of alternations. Both verb alternations and middles create (previously) non-canonical subjects and in that way act as a response to the need for subjects in the periods after the loss of V2.

6.4.1.4 Conclusion

This section has defined the properties of middles and subsequently aimed to find evidence for the claim presented in the literature that the middles are an innovation in EModE. This claim could be supported for the *it needs doing* construction, with the data from the PPCEME and PPCMBE confirming earlier findings by Visser (1963-1973) and De Smet (2013a). For the prototypical middle, however, it was a more complicated quest because of the lack of unique syntactic features of the middle, which made it difficult to select middle on the basis of a query. I found no additional data for earlier studies by Jespersen (1909-1949) and Visser (1963-1973), who gave fifteenth and sixteenth-century examples. At the same time, evidence was presented from other studies: Hundt (2007) finds an increase in use of middles in the twentieth century and Van Gelderen (2011) reports an increase in use of labile verbs throughout the history of English. In conclusion, there is suggestive evidence that the canonical middle and the *it needs doing* construction were innovations in the period after the loss of V2 and generally they fit into the broader development towards more subject strategies and more non-canonical subjects.

6.4.2 Non-agentive subjects

The final development likely to be connected to the development of post-V2 English towards a subject-initial language is the use of non-agentive subjects, or “permissive” subjects, as Los & Dreschler (2012) call them. Originally these observations were based on contrastive studies of German and English (Rohdenburg 1974), which show that English allows for a much wider range of non-agentive subjects in comparison to German. In light of the current study it is especially relevant that these observations are based on German, a modern V2 language, but Callies (2006: 117) points out that these subjects are cross-linguistically marked as well (referring to Comrie 1989: 75,79 and Blake 1987: 315, 319f.). This section first

discusses Rohdenburg's (1974) early observations, then moves on to a more recent study, from Callies (2006), which attempts to capture these differences between English and German. Finally, I address the question whether these non-agentive subjects are innovations in the history of English.

Rohdenburg (1974), as summarized in detail by Hawkins (1986), presents a large collection of examples from German and English, based on translations and original texts in both languages. Rohdenburg's main finding is that English has a wider range of non-agentive subjects than German, and Hawkins calls it the "greater semantic diversity of English basic grammatical relations" (1986: 62), stating that this diversity not only holds for subjects but also for objects. The following examples show grammatical sentences in English, their ungrammatical translated German counterparts, and the option that would be the "most natural[...]" translation (Hawkins 1986: 60).

(99) *This hotel* forbids dogs.
In diesem Hotel sind HUNDE verboten.
DOGS are forbidden in this hotel.

(100) *The latest edition of the book* has added a chapter.
In der letzten Ausgabe des Buches ist EIN KAPITEL hinzugefügt.
In the latest edition of the book A CHAPTER has been added.

(101) *This advertisement* will sell us a lot.
Mit dieser Anzeige verkaufen WIR viel.
With this advertisement WE will sell a lot.
(Rohdenburg 1974, as quoted by Hawkins 1986: 60-61)

None of the subjects in the English examples are agentive. For instance, the subject in (99) is locative, in (100) it is a theme, and in (101) it is instrumental. Even in the cases that the non-agentive subject is marginally acceptable in German, (e.g. *?Diese Zutaten backen 4 Kuchen* for *These ingredients will bake 4 cakes*), Hawkins stresses the fact that English has a clear preference over German for the non-agentive version. This means that the difference between the two languages is partly explained in terms of frequency rather than only grammaticality options; Hawkins calls it a "quantitative claim" (62). The observation that the PPs are the most natural translation is especially relevant in light of the development of English from local anchoring using presubject PPs and objects towards a subject-initial language as described in the previous chapters. The fact that German is a V2 language and may use the first position for local anchoring just like OE already points to the likelihood that these types of subjects are an innovation in English.

More recent observations about the differences between English and German are presented by Callies (2006), who carried out an experiment of translation and acceptability tests of these types of subjects with German learners of English.

Examples (102)-(104) illustrate some of the test sentences that were used in the experiment.

- (102) a. Next month sees Garbage and Lenny Kravitz in town.
b. Garbage and Lenny Kravitz will be in town next month.
- (103) a. 69 million dollars bought Michael Bloomberg the election as Mayor of New York.
b. Michael Bloomberg has bought the election as Mayor of New York with 69 million dollars.
- (104) a. A major anthrax threat in Washington has closed a second postal facility.
b. A second postal facility in Washington was closed due to a major anthrax threat.

The participants were first asked whether they thought the sentences were acceptable. In 50% of the test items, the version with an adjunct scored better than the version with a non-agentive subject (2006: 120). The reasons given by the learners for the unacceptability varies from wrong word order to strange meaning, but the most common reason given is that of “unusual subject” (122). The second task was a translation task, during which the students had to translate English sentences with non-agentive subjects into idiomatic German translations. Callies points out that the results of the translation test were not entirely reliable because some learners gave literal and “highly unidiomatic German translations” (123). Nevertheless, in half of the sentences the majority of participants changed the non-agentive subject in the English sentence into a PP in the German translation. Although the number of items and participants is small, meaning that the results need to be considered with care, Callies’s data further confirm the differences between German and English, as well as the relation to the local anchoring presubject position in V2 languages.

In light of the hypothesis that this development is related to the loss of V2, an important question is whether these non-agentive subjects represent innovations in the history of English. Several authors in fact present the claim that this is the case, and date the innovation during the ME or EModE period, but the amount of data that is available to support this claim is even more circumstantial than for the middles. Rohdenburg (1974: 12) proposes that the non-agentive subjects (his “secondary subjectivations”) are innovations in English, but only gives one example: the relevant use of *lose*, which is first used in Shakespeare’s time (so around 1600), according to the NED. At the same time he also notes that most of the non-agentive uses of verbs are not included in a dictionary. Hawkins (1986: 69) suggests that secondary subjectivations are a further development from the middle (as intransitive use of transitive verbs), and links their development to the introduction of the prototypical middle as described by Visser (1963-73). While the middles are a clear

indication of a change in semantic roles for the subject, with a non-agentive argument in a transitive clause becoming the subject of an intransitive clause, and are as such related to the non-agentive subjects, not all non-agentive subjects are part of transitivity alternations and the development of non-agentive subjects should be considered separately from the development of the middles.

The reason for the lack of diachronic data might well be the fact that these non-agentive subjects are even more difficult to find in a systematic way in a corpus. Their syntactic structure is similar to the syntactic structure of regular active sentences with an agentive subject and there is no specific coding on the subject. Moreover, the range of verbs that allows this construction is quite wide, as Rohdenburg's examples show, and there are only a few specific noun+verb combinations that often occur. Among the examples from Rohdenburg there is one verb for which the OED gives the non-agentive subject reading, *seat* (which is in fact one of Ahn & Sailor's accommodation verbs), as illustrated in examples (104)-(106).

- (104) 1828–32 Webster *Amer. Dict. Eng. Lang.*, **Seat**,..to place in a church; to assign seats to. In New England..it is customary to seat families for a year or longer time; that is, assign and appropriate seats to their use.
- (105) 1856 C. Merivale *Hist. Romans under Empire* IV. xli. 535 The first object..was to seat the greatest number of the people possible.
- (106) 1887 *Pall Mall Gaz.* 9 Sept. 2/2 Each theatre should be registered and advertised as capable of seating a specified number.

Note that not all of these examples are entirely the same sort of examples as *this theatre seats*, but the last example comes close to the intended meaning. Online searches show that these sentences frequently occur with different subjects as well.

- (107) I got me a Chrysler, it seats about 20
("Love shack", B-52)
- (108) The Gibson Amphitheatre (formerly Universal Amphitheatre) is an indoor amphitheatre located in Los Angeles, California within Universal City. It seats up to 6,189 for concerts, including 6,089 chairback seats (3,900 in the main level and 2,189 in the mezzanine).
(http://en.wikipedia.org/wiki/Gibson_Amphitheatre)
- (109) It tows, it fishes, it seats everyone.
(Advertisement for Triumph boats).

Turning to the PPCMBE, *seat* only occurs once, as "they seat themselves", but never with a non-agentive subject. As is the case for some of the earlier studies in this chapter, it seems that the PPCMBE is too small for lexical searches or for finding these somewhat marginal constructions.

Nevertheless, there are some suggestive findings. First of all, occasional early typical examples of non-agentive subjects can be found in texts not included in the PPCMBE and PPCEME, for instance in Jane Austen's *Sense and Sensibility*, from 1811.

- (110) The next morning brought Elinor a letter by the twopenny post, from Lucy herself.
(*Sense & Sensibility*, chapter 38).
- (111) The second day brought them into the cherished, or the prohibited, county of Somerset,
(*Sense & Sensibility*, chapter 42)
- (112) Poor Marianne, languid and low from the nature of her malady, and feeling herself universally ill, could no longer hope that to-morrow would find her recovered;
(*Sense & Sensibility*, chapter 43)

In all these examples, the subject is a day (*the next morning, the second day, to-morrow*), much like example (102), and these subjects are combined with transitive verbs which normally take an agentive subject.

In addition, there is some circumstantial evidence that the tendency for non-agentive subjects is a property that English developed over time. This evidence is based on studies on animacy in English in different stages, a notion which is closely related to agentivity, illustrated by the fact that all non-agentive subjects presented in this section are also inanimate. The first piece of evidence comes from Kilpiö's (1989) study of the passive in *Cura Pastoralis* and *Bede*, comparing these OE texts to their Latin sources. Kilpiö presents a study of passives in *Bede* and *Cura Pastoralis* and describes a number of changes from passive to active and the other way around, and investigates factors contributing to it. He finds that one of the factors influencing the choice of a passive or active sentence in OE is the preference in OE for animate subjects, leading to translations of passives in Latin to actives in OE (1989: 203-228; 238-240). What this shows is that OE had a preference for animate subjects, which is clearly still much stronger than it is in PDE. Sometime in the intervening periods, this must have changed.

The second piece of evidence comes from Komen et al. (2014), who compare the animacy of subjects in one text from OE to two late Modern English texts and find an increase in inanimate subjects. Their study is based on a semi-automatic annotation of animacy in the texts, with the automatic part based on the syntactic and morphological information that is already available in the annotated Penn-Helsinki corpora, such as number or person. Rather than looking at subjects individually, they identify all referents in the text, determine their animacy and subsequently consider whether they function as subjects at some point in the text. What they find is that the percentage of inanimate referents functioning at least

once as subject increases from 43.7% in the OE text to 54.7% in the Late Modern English text. While Komen et al.'s data are only based on 3 texts, they provide an interesting first indication that animacy of subjects may have changed in the history of English, again pointing towards a wider range of semantic roles for the subject.

In conclusion, the synchronic data show quite clearly a tolerance in PDE for non-agentive and inanimate subjects, and there is some circumstantial historical evidence that this is a property of subjects that English has developed over the course of the centuries. It is clear that more detailed studies are necessary to confirm this development in more detail and provide a time frame for this development, specifically with respect to the loss of V2 and its consequences.

6.4.3 Summary and Conclusions

This section has investigated the development of middles and non-agentive subjects in the history of English. The loss of V2 as described in Chapter 5 provides an explanation as to why the subject was under pressure to extend its semantic possibilities through middles and non-agentive subjects: with the subject as the most important element in clause-initial position, there is a greater pressure on the subject, in terms of the functions it has to perform. As a result, strategies to move NPs to subject position become more important, and middles and non-agentive subject constructions provide such a strategy.

The middle, a construction with an active form but a passive interpretation, seems to be an EModE innovation, or its use has increased since the EModE period. One type of middle, *it needs doing*, was traceable in the corpora and indeed seemed to be an EModE innovation. The canonical middle, *bureaucrats bribe easily*, was more difficult to trace, with no examples in the PPCEME and PPCMBE. However, there is some evidence in the literature that these middles are also innovations: Hundt (2007) describes an increase in the twentieth century, and Van Gelderen (2011) reports an overall increase in the use labile (so alternating) verbs in the history of English. The non-agentive subjects that were discussed have fewer characteristics that make them stand out syntactically, which makes them more difficult to trace in the corpora. Indeed, the main observation on these non-agentive subjects comes from contrastive studies of English with German, and clearly shows the greater tolerance in PDE for non-agentive and inanimate subjects. The historical evidence consisted only of circumstantial evidence concerning a possible increase in this tolerance between OE and late Modern or PDE. As with the middles, a more thorough study is needed to confirm these suggestive findings and describe the development in detail.

Section 6.5 Conclusion

After a study of the loss of V2 and the related (information-structural) developments in the previous chapter, this chapter focused on the consequences of the loss of V2 itself, as well as the syntactic and information-structural developments described in

the previous chapter. In particular, this chapter investigated subject-creating constructions—passives, middles and non-agentive subjects—in the period after the loss of V2, starting from the assumption that these strategies provided a response to the greater need for subjects which followed the loss of local anchoring and the increase in subject-initial clauses. In addition, these constructions function as information-rearrangers, another function which is relevant in English after the loss of alternative information-rearranging constructions, as described in the previous chapter. In this chapter, I showed that, although the data vary in their robustness, there is evidence for an increase in the use of all these constructions after the loss of V2.

Section 6.2 first addressed the question of a possible increase in the use of passives after the ME period. The data showed an increase both in use of all passives with respect to active transitive clauses as well as in the use of long passives, most clearly in the Modern British period, which is slightly later than previously proposed in the literature. This section also showed that passive subjects information-structurally fit into the mould of active subjects, showing similar behaviour in terms of information status and discourse-linking properties. Finally, the long passive by now had become a fixed syntactic construction, with a standardized way of expressing the agent and a standard word order, which in turn allowed for a new function of the long passive, namely that of topic introduction.

Section 6.3 investigated the development of the new passives after their introduction in the fourteenth century. Where in the ME period there was a clear connection between the three passives in that they were introduced in roughly the same period, the developments diverge after the ME period. What the three passives have in common is that the first subperiod in the EModE sees an increase, i.e. a further establishment in all three passives. After that, the prepositional passive slowly continues to become more frequent, while the recipient passive only seems to become really established towards the end of the MBE period. The study of the ECM passive, interestingly, showed that the active version of the ECM was still quite frequent in most of the Early Modern period and only became less frequent in the later periods, while the passive version remained stable throughout these periods and slowly acquired the evidential reading it is mostly associated with today. No longer do the developments of these passives suggest a single change in the language which can be related to the loss of V2; the link to the need for subject strategies seems to exist mostly in the general subject-creating and information-rearranging functions that are relevant for all passives.

Section 6.4, finally, focused on two final subject-creating strategies: middles and non-agentive subjects. Like the passive, these strategies create non-canonical subjects, but in contrast to the passive, they are less fixed syntactic constructions. Middles use transitive verbs intransitively, but crucially with a passive-like interpretation, while non-agentive subjects occur in regular transitive subjects, but with an unexpected pairing of verb and subject. Most observations on these constructions come from PDE studies and the diachronic data is limited. Even so, I

discussed and presented some diachronic evidence that hints at these constructions being an innovation in EModE and reflecting an overall development from OE to PDE.

In conclusion, the various studies have illustrated some possible responses to the need for subjects and information-rearranging strategies after the loss of V2. Passives, middles and non-agentive subjects all create subjects which are non-canonical, in the sense that they do not use subjects which have a prototypical thematic role of agent. While some of the studies in this chapter clearly showed the limits of the corpora, taken together they present an interesting glimpse of what may have followed the loss of V2: not just a change in the position of the subject, but a response in terms of a changing use of particular constructions which make use of the available syntactic possibilities in such a way that an information-structural need—e.g. for an unmarked discourse-old subject—can also be expressed. These studies, then, further confirm earlier observations in the literature that speakers make use of the syntactic means that are available to express the information-structural needs which are likely to be constant throughout all stages of the language.

7. Conclusion

This thesis has investigated the interaction between syntax and information structure through a study of the loss of verb second (V2) and developments in the passive in the history of English. The starting point for this investigation is provided by observations from the literature about the loss of V2, which represents a major syntactic change in the history of the language, and a series of developments which suggest that there is something like a 'rise' of the passive in English, most importantly manifested in the introduction of three new types of passives towards the end of the Middle English (ME) period. Individually, both developments call for investigation because they represent rare types of changes cross-linguistically. Many V2 languages, like present-day German and Dutch, are stable and do not change their basic word order, which raises the question why English would have undergone this change. In comparison, the new passives are cross-linguistically rare, which makes the fact that they were introduced in English even more unusual. What makes these developments stand out even more, however, is that they take place roughly simultaneously, and that the passive and V2 seem to share an information-structural function: the syntax of both the passive and the OE V2-system allows for a rearranging of arguments in the clause, which in turn can be used for information-structural purposes, such as creating a given-before-new order of information and placing given information in initial position. The shared information-structural function, combined with the syntactic changes taking place at roughly the same time, make the passive and the loss of V2 particularly suitable for investigating the

interaction between syntax and information structure. The studies in this thesis provide evidence for a complex interplay of syntactic and information-structural factors: a range of word order options interact with several information-structural functions. Syntax has a great impact on information structure because it determines the word order options available in the language. Information structure, in turn, does not influence syntax to the same extent, in the sense that it does not influence what options are available, but it has a major influence on syntactic choice.

In order to provide an answer to the main questions posed in this thesis, Section 7.1 summarizes the conclusions per chapter. Section 7.2 provides answers to the main questions, outlining the main contributions of this thesis to the discussions in the existing literature. Section 7.3, finally, addresses certain limitations of the current study and provides suggestions for further research.

Section 7.1 Summary of findings per chapter

Let us start by summarizing the findings of the individual chapters. **Chapter 2** explored in more detail the background of the proposal that the loss of V2 and the rise of the passive are connected, based on the idea that V2 and the passive share an information-structural function. Two types of V2 existed in OE: one in which the finite verb is always in second position, in C, following operators and *þa/þonne* in Spec,CP; and a second type in which the verb only moves to F, following non-operators in Spec,CP (Haerberli 2000, Fischer et al. 2000). In these clauses, the subject occurs either after the finite verb, as in (1), or after it, as in (2).

- (1) *Mid þusend searocræftum wolde se swicola deofol þone halgan wer on sume wisan beswican,*
'With a thousand wily arts did the treacherous devil strive in some way to deceive the holy man,'
(coelive,ÆLS_[Martin]:710.6427)

- (2) *ON sumere tide Martinus stah to anre upflora,*
'On one occasion Martin was mounting to an upper floor,'
(coelive,ÆLS_[Martin]:601.6358)

In information-structural terms, there are two important aspects of this second type of V2. First, the presubject position can be used for unmarked, discourse-old information, and it is used by elements of various syntactic categories—objects, PPs, adverbs (see Los 2005, 2009, 2012). Second, the position of the subject is largely determined by its information status: old subjects are likely to occur before the finite verb, while new subjects are likely to follow it (see Bech 2001, Van Kemenade & Westergaard 2012). Both properties have the effect that there is variation in word order, which can be used by speakers for information-structural purposes, both to rearrange arguments to create an unmarked given-before-new order of information,

and to place given information in initial position, the preferred position for given information. In turn, the two crucial properties of the passive that make it relevant for information-structural purposes are its rearranging of the active subject and object, and the fact that it places a non-agent in subject position; it promotes an element to subject status. Again, these options can be used for information-structural purposes: rearranging elements according to a given-before-new principle, and placing given information in initial position in the clause.

Building on proposals from Los (2005, 2009, 2012) and Seoane (2006), the shared information-structural potential between the passive and V2 provides the starting point for an exploration of the interaction between syntax and information structure. Los (2009, 2012) proposes that when the V2 system was lost in the ME period, this was not just a syntactic change, but the presubject position became a marked theme with the result that the subject became more important as the only remaining unmarked theme. This in turn led to a greater need for strategies to create subjects, such as the new passives. Seoane (2006) proposes that the passive provides an answer to changing needs in the language, both as a subject creator and as an information rearranger. These two proposals present two different scenarios: in the first scenario, the loss of V2 led to the introduction of new constructions, specifically the ECM, prepositional and recipient passive; in the second scenario, the loss of V2 influenced the use of the passive a less direct way, leading to an increase in the use of passives, including the three new passives, and possible other strategies to place arguments in subject position.

The final part of Chapter 2 reviewed how information-structural functions have previously been defined for a range of constructions, either as absolute constraints on the information status of a particular element, as relative constraints, or with information status as one factor working together with factors such as weight. I also addressed the question of how to understand and annotate givenness and concluded that cognitive accessibility and discourse behaviour should be considered separately because they are sometimes independent factors determining the use of a construction.

Chapter 3 investigated the introduction of the three new passives in the fourteenth century and aimed to establish the time and reason for their introduction, in order to determine whether the loss of V2 led directly to the introduction of the passives. I showed that their introduction is even closer in time than previously proposed in the literature and that, crucially, they are all in use, although marginally, before the loss of V2. In Section 3.2 I argued that the introduction of the prepositional passive, which authors previously proposed to have taken place in the thirteenth century, actually takes place slightly later, with the first unambiguous example dating from the 1330s. I provided additional evidence for the process of reanalysis of verb and preposition as the key to the introduction of the prepositional passive. I explained the introduction in terms of minimal alternations, where the new passives are first introduced in a context where they are least noticeable—in relative clauses, where stranding was already possible.

The studies of the recipient passive in Section 3.3 provided no earlier examples than previously found by Allen (1995), from the late fourteenth century. I argued that as long as the dative-fronted passive is still in use (up to roughly 1400), the fact that the recipients could be placed in initial position as datives blocked nominative recipients. Once fronting in general declines, including fronting of datives, the recipient passive becomes possible.

Finally, the study of the ECM construction in Section 3.4 added examples to the existing collection in the literature, but the corpus contained no earlier examples, and crucially did not contain more active ECMs than previously described in the literature, in turn confirming the dominance of the passive version of this construction. I argued that this dominance is due to the closely-related passive options that were already available in the language, making the passive unmarked, while the active stands out syntactically because of its unusual structural configuration. Crucially, what these three studies show is that the introduction of each new passive is caused by specific small-scale changes: the reanalysis of verb and preposition in the case of the *for* prepositional passives; the loss of the dative-fronted passive in the case of the recipient passive; and influence of Latin and the existence of related passives in the case of the ECM passive. There is no clear evidence of information-structural factors playing a role in their introduction.

Chapter 4 investigated the information-structural aspects of the V2 system in OE and aimed to establish the way syntactic and information-structural factors interacted in OE. I showed that various constructions are in OE available for similar purposes of achieving a given-before-new order of information: the passive, object fronting, and PP preposing. In addition, the subject of the passive, the fronted object and the initial PP all performed a function of local anchoring by allowing a given, discourse-old element to occur in initial position in the clause.

Section 4.2 showed how the passive is syntactically not completely standardized in OE: both the auxiliaries and agent phrase preposition are not yet fixed. What is in place, however, is the function of the long passive in terms of rearranging information, which becomes clear from the results that almost ninety percent of the long passives obey a given-before-new constraint. I found that late subjects (cf. PDE *Into the room came an elephant*), for which passive clauses are one of the possible verb types, provide an option for speakers to present new subjects in end position in the clause, i.e. in focus position.

Section 4.3 showed that object fronting in OE obeys a similar type of given-before-new constraint—the object needs to be more given than the subject—but the percentage is slightly lower than that for passives, around 75%. What also makes object fronting different from the passive is that most of the exceptions to the given-before-new constraint are instances of contrastive or focused clause-initial objects. This was further confirmed by a study of clause-initial objects, which showed both contrastive or focused objects in that position, but the majority of objects in clause-initial position (quite different from PDE) contain old information and are information-structurally unmarked.

In Section 4.4, it was shown that, like the clause-initial objects, clause-initial PPs in OE mostly contain discourse-old information and function as unmarked discourse-links, i.e. local anchors. Like objects in initial position, clause-initial PPs can also be marked or contrastive. For all three constructions, then, the studies provided evidence for word order variation being used for the information-structural purpose of placing given information before new information. This overlap is most clearly seen in the case of object fronting and passive, as the constructions can perform the same function and target the same arguments and can thus be truly said to be in competition.

Chapter 5 investigated the remnants of the V2-system in the EModE and MBE periods and aimed to achieve more insight into the loss of inversion in PP-initial clauses and into the discourse behaviour of these initial PPs. It showed that what is relevant in the change described as the loss of V2 is not only the change in the position of the finite verb and subject, but also a development in the use and function of the subject and the element in initial position. Section 5.2 showed that the position of the subject changed: during the EModE period, inversion is lost almost completely, with the exception of clauses with unaccusative verbs. The remnants that remain from the beginning of the MBE period already resemble the remnants in PDE in the locative inversion pattern. Moreover, Section 5.3 showed that during EModE, but even more convincingly during the MBE period, English develops into a predominantly subject-initial language. This has the consequence that the presubject position becomes less frequently used, and the distribution of the types of elements in initial position shifts. Crucially, the information-structural character of the elements in initial position also changes: PPs become more often frame setters and contrastive elements, and object fronting is almost completely restricted to contrastive objects. In short, this chapter identified in more detail the changes that are related to, or are consequences of, the syntactic development that is the loss of V2, most importantly manifested in the increasingly marked nature of the presubject position, and the emergence of the subject as the only remaining unmarked theme.

Chapter 6 returned to the topic of the passive and other subject-creating strategies, and traced the developments in these constructions in the periods after ME, up to the beginning of the twentieth century, with respect to the changes in the language during and after the loss of V2. I showed that these strategies can be seen as another response to the need for subject strategies after the loss of V2 because of their subject-creating and information-rearranging properties. In Section 6.2, I provided further evidence for the previously described increase in the use of passives with respect to actives; although there are many differences between contexts and between subperiods, a convincing overall increase takes place from EModE to MBE. A similar increase takes place specifically for long passives, which are relevant not only for their subject-creating properties but also for their information-rearranging function. In Section 6.3, I traced the development of the three new passives towards becoming a more established part of the English language: they all become gradually more frequent in the language and develop into productive

patterns, although they are all limited somewhat by lexical restrictions. Finally, in Section 6.4, I explored middles and non-agentive subjects and provided suggestive evidence from a range of sources that these are innovative strategies in the period after the loss of V2 and that their use increased in these periods. All these constructions represent strategies which fulfil the need for subject strategies in that they provide non-canonical ways of placing an argument in subject position in the clause.

Section 7.2 Answers to the main questions

The main aim of this thesis is to investigate the interaction between syntax and information structure, and specifically (i) whether there is evidence of a two-way interaction at any stage of the language; and (ii) whether the interaction between syntactic and information-structural changes throughout the centuries, when we see the syntactic changes of the loss of V2 and introduction of new passives. Most importantly, did syntactic changes limit the information-structural means to such an extent that new syntactic constructions were introduced to express the information-structural functions?

7.2.1 The interaction between syntactic and information-structural factors in Old English

The first question to be answered is what role information-structural factors play in the word order variation in OE (450-1150) and to what extent these factors can be seen as the main motivation for the use of particular constructions. The answer to this, based on the corpus studies from Chapter 4, is that information structure plays an important role in OE V2, as can be seen in the fact that information-structural factors determine to a large extent variation in subject placement, as well as the choice for particular constructions. First, the presubject position is used for the specific information-structural function of local anchoring, as we have seen in object-initial and PP-initial clauses: these types of clauses allow for a discourse-old referent to occur in the information-structurally unmarked initial position. Second, passives, object fronting, and to a lesser extent PP preposing, are all available for the same information-structural function of reordering elements in the clause according to a given-before-new order. This overlap in function is clearest for passives and object fronting because they target the same arguments in the clause for rearranging (the underlying object and subject of an active). Third, the three subject positions that are available in non-operator-initial clauses, where V-to-F movement takes place, are differentiated in terms of the information status of the subjects occurring in those position: the studies in Chapter 4 show that there is a general tendency for old subjects to occur before the finite verb and for new subjects to follow the finite verb, confirming earlier proposals in the literature (e.g. Bech 2001, Van Kemenade & Milićev 2012 and Hinterhölzl & Petrova 2010). The studies in this thesis update these findings by differentiating between V2 contexts, which is not

always consistently done in the literature (V-to-F vs V-to-C, pronominal vs nominal subject), and by using finer-grained information status distinctions. The studies show that while the general tendency is clear, it is not a strict rule, witnessed by the many counterexamples. In a similar way, the late subject position provides another option for placing new information late in the clause, already prefiguring the specific introducing function of locative inversion in later periods.

However, while information-structural factors clearly influence the use of particular constructions, they do not regulate the system to such an extent that particular positions or constructions can be said to be strictly driven by the need to express an information-structural function. Crucially, the different subject positions do not correlate with specific information statuses. This means that theoretical accounts should provide room for some variability, something which not all of the current proposals on OE syntax and information structure allow for (e.g. Hinterhölzl & Petrova 2010 and Biberauer & Van Kemenade 2011). Although optionality in syntax may not be an ideal conclusion, it does provide the most elegant account of the interaction between syntactic and information-structural principles, while also leaving room to consider related factors, such as weight.

7.2.2 Changes in the interaction between syntax and information structure after the loss of V2

The second question with respect to the interaction between syntax and information structure is how the role of information-structural factors changes when V2 as a syntactic system is lost. The Middle English period (1150-1500) sees the beginning of the loss of V2, in the decline of subject-verb inversion in non-operator contexts, from 1400 onwards. The change takes place at such a speed that V2 is lost for the greater part by 1500, although it is clear that the loss of V2 does not represent a uniform development: in some contexts—most notably after PPs and with unaccusative verbs—inversion is still relatively common by the end of the ME period. The loss of V2 seems to be as much an information-structural development as a syntactic change (cf. Los 2009, 2012). The developments in the subject and presubject position during the transition from a V2 language to an SVO language, as described in Chapter 5, show a combination of syntactic changes, changes in the information-structural character of positions, and changes in frequency in the use of particular positions. In the first place, the position of the subject with respect to the finite verb changes, as seen in the study of inversion in PP-initial clauses in Chapter 5, which added to the existing body of literature on the loss of V2 that has previously focused on clause-initial adverbs and objects (e.g. Haeberli 2002, Van Kemenade & Westergaard 2012, Van Kemenade 2012). Inversion of the old V2-type occurs for a longer time in PP-initial clauses, until the E1 period, while late subjects still remain possible in PDE, although limited by the type of PP, type of verb and information status of PP and subject. Inversion after PPs, which is still relatively common at the beginning of the Early Modern English period—including both old V2-type inversion examples and locative inversion—shows a sharp decline by the end of the sixteenth

century. The loss of one of the subject positions means that there was less room for the given-before-new principle to order subjects according to their information status or anaphoric behaviour.

Second, during the same period that the position of the subject becomes more fixed before the finite verb, the use of the clause-initial position changes. The growing need for subject strategies in the language becomes clear through the dominance of subjects in initial position in the EModE and MBE periods. This dominance of the subject has an influence on the use of the initial position: object fronting is lost, in two ways: objects become infrequent in initial position, and they lose their local anchoring function, which also means a loss of a rearranging construction for given-before-new order. PP preposing is not lost but changes in character: PPs become increasingly marked, although some local anchoring remains; crucially, however, local anchoring is no longer the most important function of clause-initial PPs. The EModE and MBE periods show that what is relevant in the change described as the loss of V2 is not only the change in the position of the finite verb and subject, but also a development in the use and function of the subject and the element in initial position, confirming Los's (2012: 21) statement that V2 was "much more than a gradual falling off of the frequency of verb movement" (see also Van Kemenade 2012). With the loss of V2, English becomes an overwhelmingly subject-initial language, leading to a situation in which other elements in first position become marked; even more so because of their increasingly marked information-structural properties of frame setting and providing contrast. In sum, what we see is that changes in the syntax limit the options that are available for speakers to express particular information-structural functions.

7.2.3 Responses to the loss of V2

The third question is whether the developments in the passive and other subject-creating strategies can be ascribed to the changes in interaction between syntax and information structure and more specifically, to the limitations exerted by syntax on information-structural freedom. I have explored three possible consequences of the greater importance of the subject: (i) the introduction of new passives in the fourteenth century; (ii) an increase in use of the passive; and (iii) a rise of alternative subject-creating strategies.

As for the introduction of new passives, I have shown in chapter 3 that the loss of V2 in terms of information structure or more precisely, the changes in the use of the presubject and subject position, did not lead to the introduction of the three new passives: the passives were in use before the loss of V2 and their introduction can be explained with reference to specific syntactic developments or contexts. The passives were introduced by the end of the ME period, with the dates of first attested examples lying close together: the prepositional passive in around 1350, the recipient passive around 1375, and the ECM passive around 1390. Despite the suggestive timing of the three introductions, the passives are neither structurally linked to one another, nor more generally to the effects of the loss of V2. In the

period that the consequences of the loss of V2 kick in, and the need for subject strategies becomes greater because of the increased importance of the subject as the initial element in the sentence, the new passives become more established constructions, as shown in an increase in frequency and productivity. Their function as subject creators and information rearrangers makes them particularly suited constructions as a subject strategy and a response to the increasing restrictions on word order freedom.

In this respect, the new passives follow the general development of all passives, which show an increase in frequency throughout the EModE and MBE periods. By this time, the passive is a fixed syntactic construction, with one way of expressing the agent, through the *by*-phrase, and an increasingly fixed position for this *by*-phrase after the passive participle, in postverbal (i.e. focus) position in the clause. First, the passives are relevant as responses because of their subject-creating properties. The studies in Chapter 6 provide evidence for an increase of passives with respect to actives in EModE, confirming Seoane's (1999, 2000, 2006) data. The increase is most pronounced, however, if the EModE and MBE corpora are considered in their entirety to each other, while it is less clear in which of the subperiods this development exactly takes place. Second, passives become more strongly information-rearranging constructions after the ME period—as their syntax becomes fixed and they remain as one of the few means available to do this type of rearranging. This in turn is likely to be one of the reasons for the increase in long passives, which provide this option of rearranging arguments most clearly.

A third consequence of the loss of V2 was found in other unusual subject strategies—middles and non-agentive subjects—which are minor innovations in English and serve the same function as the passive, positioning a non-agentive argument in grammatical subject position. While the introduction of these patterns proved somewhat difficult to trace, these strategies are productive patterns in Modern English and there is evidence that they have become more frequent over the last two centuries (e.g. Visser 1963-1973, Hundt 2007, De Smet 2013). This is the period that the information-structural consequences of V2 have truly taken their effect, i.e. when English has become a predominantly subject-initial language and subjects remain the only possibility to express unmarked themes. The development of middles and non-agentive subjects fits into the overall story of the need for subject strategies and increased use of non-canonical subject creators and information-rearrangers.

7.2.4 Conclusion: a complex interplay of syntactic and information-structural factors

The overall conclusion regarding the relation between the loss of V2 and the 'rise' of the passive is that the loss of V2 and the information-structural consequences of this loss did not lead to syntactic changes in the form of new constructions, but did boost the use of certain existing constructions from the sixteenth and seventeenth century onwards, when the information-structural aspects of the loss of V2 had been

established and the subject had become the only unmarked theme. This means that information structure can, on the basis of this study, not be shown to influence the syntactic options that are available and, in turn, that the current study does not provide evidence for an equal two-way interaction between syntax and information structure. Instead, the studies provide evidence for the view that information structure makes use of the options that the syntactic system makes available (cf. Van Kemenade & Westergaard 2012: 105). When the syntax changes, the distribution in terms of frequency of the remaining options to express a particular function shifts, because the same information-structural needs still must be met. This is most clearly seen in the subject positions: as long as there are multiple positions, information structure plays an important role in determining the variation but as soon as these positions are lost, speakers look for other existing means to express some of these functions, for example introducing new topics through locative inversion and passives. The same interaction is also seen in the changes in the number of constructions that can achieve given-before-new order: from a range of constructions in OE to only the passive in PDE; the increase in the use of the passive following this development is more than suggestive.

At the same time, the studies have also shown that information structure has a considerable influence on the language and the changes that we have seen in the Left Periphery—mostly information-structural—in fact determine the shape of the language to a large extent: while all syntactic positions remain available, the character of the presubject position changes dramatically, accompanied by changes in the frequencies of elements occurring in this position. None of these changes are syntactic, but they do have a far-reaching effect on the use of the initial position. It is important to recognize that these information-structural functions can be of a different type: in some cases conditions on information status seem absolute, in other cases it is a matter of relative constraints. The studies provide evidence that information-structural needs are more stable than the syntax in a language: the principles of given-before-new order of information and marked/unmarked sentence beginnings are important in all periods, while the syntactic means change (cf. Lambrecht 1994: 29). Finally, although there is no evidence of syntactic changes as the result of the loss V2, the changes in the use of particular constructions—a continuing decrease in frequencies leading to limited productivity—may lead to syntactic changes in the long run. If we return to the Minimalist principles, it is clear that information structure is part of the language use, and the current study has not provided evidence for information structure as part of the syntactic system of the language (cf. Chomsky 1995). As Minimalists, then, we should not want to incorporate information structure in the syntactic derivation, but work with a model which allows for a level of optionality and where the choice between options is determined by the speakers on the basis of information-structural needs. An insight into these syntactic possibilities, in combination with separate studies on these information-structural needs, provides a valuable perspective on the changes that

have taken place in English, but the value of describing such an interaction is not limited to this particular case study.

Section 7.3 Suggestions for further research

This final section outlines some unresolved issues and provides suggestions for further research. While the corpus studies have provided an interesting insight into the interaction between syntactic and information-structural factors in the particular case of passives and the loss of V2, some questions remain unresolved. One obvious limitation is that the current study has only considered English, while the interaction between syntax and information structure can only be fully understood when all languages are taken into account. An example of a question that has not been sufficiently answered is what the influence of genre is on the use of the passive. In addition I have focused on comparing the passive with V2-type constructions only, but there are more constructions that share a function with the passive, including impersonalization strategies. In this section, I highlight three of these issues, which in turn provide suggestions for further research: (i) the annotation of information status; (ii) the challenges of corpus studies; and (iii) future developments in English.

One of the principal decisions on the basis of the literature on annotation of information structure in both PDE and OE was to use a four-way distinction of information statuses in combination with separate ways of measuring anaphoricity. The corpus studies have underlined the relevance of separating anaphoricity from cognitive accessibility in describing the function of a particular construction. For instance, the study of topic introduction as a function of the long passive illustrated how it is not information status but anaphoricity which seems to be the deciding factor in the use of the construction. The choice of four information status categories—more detailed than most studies on related topics, which mostly use binary distinctions—yielded more mixed results. In some cases, the finer-grained distinctions provide a better understanding of the relevant contrast, for instance in those cases where the four statuses behave in terms of a hierarchy. In other cases, it was not at all clear that these four categories really constitute a hierarchy and that the categories of ACCESSIBLE and ANCHORED really added something to a binary distinction. Crucially, it is not evident whether it really makes a difference in what way a referent becomes accessible, as long as it is indeed accessible to the hearer. In that sense, while the different categories and anaphoricity provide further insight into the concept of givenness, this may not always translate into a functional difference in the use of a construction. In addition, using finer-grained distinctions also means that the choices may be more difficult to make; as documented, sometimes the status of elements in a binary given/new system is not even clear. For example, should generics and world knowledge be grouped with accessible or discourse-old referents? And is the link to the discourse or the hearer in the case of different types of inferables strong enough to make the referent behave like a given referent? Research on behaviour of all subcategories in different information-

structural contexts (for an example, see Taylor & Pintzuk 2014) is needed to fully answer the question which subcategories behave in similar ways and what type of higher-level classification is feasible.

Throughout this thesis I have combined large quantitative studies with small-scale qualitative studies, and I believe this has provided an interesting insight into the changes that take place, although both types of approaches have their limitations. First, adding qualitative studies to quantitative studies decreases the possibility that the overall numbers do not represent the changes that we are interested in, since large studies have the problem that there may be issues with the annotation of the corpus; a query for the entire corpus necessarily relies on the annotation that is available in the corpus. Second, adding quantitative studies to qualitative studies solves the problem of generalizability of qualitative studies. By singling out relevant distinctions from the quantitative studies and using these for a larger-scale study, it is possible to provide a more general picture of what has been established as a qualitative development (such as for the presence of anaphoric elements in clause-initial PPs). Despite this, there are certain gaps in the data that I have not been able to address. The most important issues in this respect are possible changes in animacy and discourse-linking properties of clause-initial subjects, possible differences between discourse-linking properties of subjects in active versus passive sentences, and the development of middles and the non-agentive subjects. For most of these developments, it has not been possible to trace the development throughout the subperiods. More detailed corpus work on a variety of corpora in combination with more extensive manual analyses (cf. Hundt 2007) is needed to fully understand these developments and to provide a time frame for them.

A final area of interest in relation to the topics discussed in this thesis that I would like to address here is the question of further changes in the language after the periods that were investigated, and indeed, after the present day. The developments and changes studied in this thesis do not, for the majority, represent single syntactic changes, but are more subtle developments that take place on a level of language use, rather than a syntactic level—with the exception of the introduction of the new passives. The corpora that have been used only include texts up to 1914, while the study of the middles has shown that important changes still take place in the twentieth century. Not only does this mean that the developments described in this thesis are likely to continue, but it also opens the possibility for a change of direction in these developments under the influence of new influences in the language. Of particular interest in this respect is the fact that the community of speakers of English has changed dramatically in this last century, with English being increasingly used as a global language or *lingua franca* (see Crystal 2003). This raises two important questions. First, with respect to the use of the clause-initial position, the question is whether learners of English use the presubject and subject position differently under influence of their first language (L1), and whether English itself, or certain varieties of English, change under the influence of these languages. It is of

course especially interesting to see the effect on English of an L1 language which has V2 (cf. Los, Verheijen & De Haan 2013). Second, with respect to the particular constructions discussed in this thesis—the three passives, locative inversion, middles—the question is how easily these constructions are acquired by L2 speakers of English. Indeed, they may represent features that are lost in English as a *Lingua Franca*, or their use may change in new varieties of English as speakers use it in a different way. An insight into these developments could provide further support for the idea that information-structural principles are universal and shed more light on the interaction between syntax and information structure.

Bibliography

- Aboh, E.O. 2010. Information Structuring begins with the Numeration. *Iberia: An International Journal of Theoretical Linguistics* 2(1): 12–42.
- Ackema, P. & Schoorlemmer, M. 1995. Middles and nonmovement. *Linguistic Inquiry* 26(2): 173–197.
- Ahn, B. & Sailor, G. To Appear. The Emerging Middle Class, In Baglini, R., Grinsell, T., Keane, J. Roth Singerman, A. & Thomas, J. (eds.), *Proceedings of The 46th Annual Meeting of the Chicago Linguistic Society* (CLS 46). Chicago: The Chicago Linguistic Society.
- Alexiadou, A. & Anagnostopoulou, E. 2004. Voice Morphology in the Causative-Inchoative Alternation: Evidence for a Non-Unified Structural Analysis of Unaccusatives. In Alexiadou, A., Anagnostopoulou, E. & Everaert, M. (eds.), *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface*, 115–136. Oxford: Oxford University Press.
- Allen, C.L. 1980. Movement and deletion in Old English. *Linguistic Inquiry* 11: 261–325.
- Allen, C.L. 1995. *Case marking and reanalysis: grammatical relations from Old to Early Modern English*. Oxford: Oxford University Press.
- Ariel, M. 1988. Referring and accessibility. *Journal of Linguistics* 24: 67–87.
- Baker, M. 1988. *Incorporation: a theory of grammatical function changing*. Chicago: Chicago University Press.
- Baker, M., Johnson, K., & Roberts, I. 1989. Passive Arguments Raised. *Linguistic Inquiry* 20(2): 219–251.
- Bately, J. 1980 (ed.). *The Old English Orosius* (EETS Supplementary Series 6). London: Oxford University Press.
- Bean, M. C. 1983. *The development of word order patterns in Old English*. London: Croom Helm.
- Bech, K. 2001. *Word Order Patterns in Old and Middle English: a syntactic and pragmatic study*. PhD dissertation, University of Bergen.
- Bech, K. In prep. Old 'truths', new corpora: Revisiting the word order of conjunct clauses in Old English.
- Behaghel, O. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25: 110–142.
- Behaghel, O. 1932. *Deutsche Syntax: Eine geschichtliche Darstellung*, Vol. 4: *Wortstellung. Periodenbau*. Heidelberg: Carl Winter.
- Bergen, L. van. 2000. The indefinite pronoun *man*: "nominal" or "pronominal"? In Bermudez-Otero, R., Denison, D., McCully, C.B. & Hogg, R.M. (eds.), *Generative Theory and Corpus Studies: A dialogue from 10 ICEHL* (Topics in English Linguistics 31), 103–122. Berlin: De Gruyter.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Biberauer, T. & Roberts, I. 2005. Changing EPP parameters in the history of English: accounting for variation and change. *English Language and Linguistics* 9(1): 5–46
- Biberauer, T. & van Kemenade, A.M.C. 2011. Subject positions and information-structural diversification in the history of English. *Catalan Journal of Linguistics* 10: 17–69.
- Bickerton, D. 1973. The nature of a creole continuum. *Language* 49: 640–669.
- Bickerton, D. 1975. *Dynamics of a creole system*. London: Cambridge University Press.

- Birner, B., & Ward, G. 1998. *Information Status and Canonical Word Order in English*. Amsterdam: John Benjamins.
- Birner, B. 1996. Form and function in English *by*-phrase passives. In *Proceedings of the 32nd Annual Meeting*, Chicago Linguistic Society, 23–31 .
- Birner, B. 2006. Inferential relations and noncanonical word order. In Birner, B. & Ward, G. (eds.), *Drawing the boundaries of meaning: Neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*, 31–51. Amsterdam: John Benjamins.
- Blake, B. 1987. English and German: Two languages two thousand years apart. *Multilingual* 6: 309–323.
- Bock, H. 1931. Studien zum präpositionalen Infinitiv und Akkusativ mit dem *to*-Infinitiv. *Anglia* 55: 114–249.
- Bohnacker, U. & Rosén, C.. 2007. How to start a V2 declarative clause: Transfer of syntax vs. information structure in L2 German. In Anderssen, M. & Westergaard, M. (eds.). *Nordlyd* 34(3): *Papers from the Language Acquisition Workshop SCL 2006*, 29–56. Tromsø: CASTL
- Bolinger, D. 1965. *Forms of English*. Cambridge, Mass. Harvard University Press.
- Bolinger, D. 1974. On the Passive in English. In Makkai, A. & Becker, V. (eds.), *The First LACUS Forum, 1974*, 57–80. Columbia, S.C.: Hornbeam.
- Bolinger, D. 1977. *Form and Meaning*. London: Longman.
- Booij, G. & Kemenade, A.M.C. van. 2003. Preverbs: an introduction. In Booij, G. and Marle, J. van (eds.), *Yearbook of Morphology 2003*, 1–11. Dordrecht: Kluwer Academic Publishers.
- Bosworth, J. & Toller, T.N. (eds). 1898. *An Anglo-Saxon Dictionary*. Oxford: Oxford University Press
- Bosworth, J. (ed). 1859. *King Alfred's Anglo-Saxon Version of the Compendious History of the World by Orosius*. London: Longman, Brown, Green and Longmans.
- Breban, T. 2012. Functional shifts and the development of English determiners. In Lopez-Couso, M., Los, B.L.J. & Meurman-Solin, A. (eds.). *Information Structure and Syntactic Change*, 271–300. Oxford: Oxford University Press.
- Breivik, L. E. & Swan, T. 1994. Initial adverbials and word order in English with special reference to the Early Modern English period. In Kastovsky, D. (ed.). *Studies in Early Modern English*, 11–43. Berlin: Mouton de Gruyter.
- Bresnan, J. & Kanerva, J.M. 1989. Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* 20: 1–50.
- Bresnan, J. 1994. Locative Inversion and the Architecture of Universal Grammar. *Language* 70: 72–131.
- Bresnan, J., Cueni, A. Nikitina, T. & Baayen, R.H. 2005. Predicting the Dative Alternation. In Bouma, G., Kraemer, I. & Zwarts, J. (eds.). *Cognitive Foundations of Interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.
- Burzio, L. 1981. *Intransitive Verbs and Italian Auxiliaries*, PhD dissertation, MIT, Cambridge, Mass.
- Burzio, L. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.
- Callies, M. 2006. Why Money Can't Buy You Anything in German: A Functional-Typological Approach to the Mapping of Semantic Roles to Syntactic Functions in SLA. In Arabski, J. (ed.). *Crosslinguistic Influences in the Second Language Lexicon*, 111–29. Clevedon: Multilingual Matters.
- Canale, M. 1978. *Word order change in Old English: base re-analysis in generative grammar*. Unpublished PhD dissertation, McGill University Montreal.
- Chafe, W.L. 1970. *Meaning and the structure of language*. Chicago: University of Chicago Press.

- Chafe, W.L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and points of view. In Li, C.N. (ed.), *Subject and Topic*, 25–55. New York: Academic Press.
- Chafe, W.L. 1987. Cognitive constraints and information flow. In Tomlin, R. (ed.), *Coherence and grounding in discourse: outcome of a Symposium*, 21–51. Amsterdam: John Benjamins.
- Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton
- Chomsky, N. 1974. The Amherst lectures. *Documents Linguistiques*, Université de Paris VII.
- Chomsky, N. 1981. *Lectures on Government and Binding* (Studies in Generative Grammar 9). Dordrecht: Foris.
- Chomsky, N. 1986. *Barriers* (Linguistic Inquiry Monograph 13). Cambridge, Mass.: MIT Press.
- Chomsky, N. 1991. Some notes on economy of derivation and representation. In Freidin, R. (ed.), *Principles and parameters of comparative syntax*, 417–454. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1995. The Minimalist Program. Cambridge, Mass.: MIT Press.
- Chomsky, N. 2001. Derivation by Phase. In Kenstowicz, M. (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, Mass.: MIT Press.
- Clark, H. and Haviland, S. 1977. Comprehension and the given-new contract. In Freedle, R. (ed.), *Discourse Production and Comprehension*, 1–40. Norwood, NJ: Ablex.
- Collins, C. 2005. A smuggling approach to the passive in English. *Syntax* 8(2): 81–120.
- Comrie, B. 1989. *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: University of Chicago Press.
- Comrie, B. 2006. Transitivity Pairs, Markedness, and Diachronic Stability. *Linguistics* 44(2): 303–318.
- Coopmans, P. 1989. Where stylistic and syntactic processes meet: Locative inversion in English. *Language* 65(4): 728–751.
- Cornelis, L. 1997. *Passive and perspective*. PhD dissertation Utrecht University. Amsterdam/Atlanta: Rodopi.
- Crystal, D. 2003. *English as a global language*. 2nd edn. Cambridge: Cambridge University Press.
- Culicover, P.W. & Levine, R.D. 2001. Stylistic inversion in English: a reconsideration. *Natural Language & Linguistic Theory* 19: 283–310.
- Culicover, P. W. 2001. Parasitic Gaps: A History. In Culicover, P.W. & Postal, P.M. (eds.), *Parasitic Gaps*, 3–68. Cambridge, Mass.: MIT Press.
- Davies, W.D. & Dubinsky, S. 2004. *The grammar of Raising and Control: A course in syntactic argumentation*. Oxford: Blackwell.
- De Smet, H. 2005. The Corpus of Late Modern English Texts (CLMET), <<http://perswww.kuleuven.be/~u0044428/clmet.htm>>
- De Smet, H. 2013a. Change through recombination: blending and analogy. *Language Sciences* 40: 80–94.
- De Smet, H. 2013b. *Spreading patterns: Diffusional change in the English system of complementation* (Oxford Studies in the History of English). New York: Oxford University Press.
- Denison, D. 1981. Aspects of the history of English group-verbs: with particular attention to the syntax of the Ormulum. Unpublished PhD dissertation, University of Oxford.
- Denison, D. 1985. Why Old English had no prepositional passive. *English Studies* 66(3): 189–204.
- Denison, D. 1993. *English Historical Syntax: Verbal Constructions*. Harlow: Longman Group.

- Denison, D. 2006. Category change and gradience in the determiner system. In Kemenade, A.M.C. van & Los, B.L.J. (eds.). *The Handbook of the History of English*, 279–304. Oxford: Blackwell
- Dik, S.C. 1978. *Functional Grammar*. Amsterdam: North-Holland Publishing Co.
- Dixon, R.M.W. 1991. *A new approach to English grammar, on semantic principles*. Oxford: Clarendon Press.
- Dobson, E.J. 1972. *The English text of the Ancrene riwle edited from B.M. Cotton ms. Cleopatra C vi*. (EETS Original Series 267). London: Oxford University Press.
- Dobson, E.J. 1976. *The origins of Ancrene Wisse*. Oxford: Clarendon.
- Downing, A. & P. Locke. 1995. *A university course in English grammar*. Hemel Hempstead: Phoenix ELT.
- Dreschler, G.A. & Hebing, R.G.A. 2011. Sentence Beginnings in Present-Day English, Present-Day Dutch and Old English. In Cummins, C., Elder, C-H., Godard, T., MacLeod, M., Schmidt, E. & Walkden, G. 2011. *Proceedings of the Sixth Cambridge Postgraduate Conference in Language Research*, 58–69. Cambridge: Cambridge Institute of Language Research.
- Dreschler, G.A. 2010. Topic introduction in the *by*-phrase: Long passives in Early Modern English. In Kampen, J. van & Nouwen, R. (eds.). *Linguistics in the Netherlands*, 59–72. Amsterdam: John Benjamins.
- Dreschler, G.A. 2014. Tracing overlap in function in historical corpora: a case study of English object fronting and passivization. In Bech, K. & Eide, K.G. (eds.). *Information Structure and Syntactic Change in Germanic and Romance Languages*, 111–139. Amsterdam: John Benjamins.
- Einenkel, E. 1887. *Streifzüge durch die mittlenglische Syntax unter besonderer Berücksichtigung der Sprache Chaucer's*. Münster: H. Schöningh.
- Embick, D. 1998. Voice systems and the syntax/morphology interface. In Harley, H. (ed.) *The Proceedings of the Penn/MIT Workshop on Aspect, Argument Structure, and Events*, MITWPL, 41–72.
- Embick, D. 2004a. Unaccusative Syntax and Verbal Alternations. In Alexiadou, A., Anagnostopoulou, E. & Everaert, M. (eds.). *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*, 137–158. Oxford: Oxford University Press.
- Embick, D. 2004b. On the Structure of Resultative Participles in English. *Linguistic Inquiry* 35(3): 355–392.
- Evert, S. 2006. How random is a corpus? The library metaphor. *Zeitschrift für Anglistik und Amerikanistik* 54(2): 177–190.
- Firbas, J. 1966a. On defining the theme in functional sentence analysis. In Daneš, F., Horálek, K., Skalička, V, Trost, P. & Vachek, J. (eds.). *Travaux Linguistiques de Prague*, Vol 1, 267–280.
- Firbas, J. 1966b. Non-thematic subjects in contemporary English. In Daneš, F., Horálek, K., Skalička, V, Trost, P. & Vachek, J. (eds.). *Travaux Linguistiques de Prague*, Vol 2, 239–256.
- Fischer, O.C.M. & Wurff, W. van der. Syntax. In Hogg, R., Denison, D. (eds.). *A History of the English Language*, 109–198. Cambridge: Cambridge University Press.
- Fischer, O.C.M. 1989. The origin and spread of the accusative and infinitive construction in English. *Folia Linguistica Historica* 8: 143–217.
- Fischer, O.C.M. 1994. The fortunes of the Latin-type accusative and infinitive construction in Dutch and English compared. In Swan, T., Mørck, E. & Jansen Webstvik, O. (eds.). *Language change and language structure: Older Germanic languages in a comparative perspective* (Trends in Linguistics: Studies and monographs 73), 91–133. Berlin: Mouton de Gruyter.

- Fischer, O.C.M., Kemenade, A.M.C. van, Koopman, W., & Wurff, W. van der. 2000. *The Syntax of Early English*. Cambridge: Cambridge University Press.
- Folli, R. & Harley, H. 2005. Flavors of v: Consuming results in Italian and English. In Slabakova, R. & Kempchinsky, P. (eds.), *Aspectual Inquiries*, 95–120. Dordrecht: Kluwer.
- Frary, L.G. 1929. Studies in the syntax of the OE passive, with special reference to the use of 'Wesan' and 'Weorðan' (Language Dissertations 5). Baltimore: Linguistic Society of America.
- Fristedt, S.L., 1953. *The Wycliffe Bible*, 2 vols (Stockholm: Almqvist & Wiksell, 1953-1969), Part I: *The Principal Problems Connected with Forshall and Madden's Edition*, Stockholm Studies in English, 4, p. 15.
- Fristedt, S.L. 1969. *The Wycliffe Bible*, 2 vols (Stockholm: Almqvist & Wiksell, 1953-1969, Part II: *The Origin of the First Revision as presented in De Salutaribus Documentis*, Stockholm Studies in English, 21, p. LVI.
- Fujita, K. 1994. Middle, ergative & passive in English: a minimalist perspective. In *The morphology-syntax connection, volume 22 of MITWPL*, 71–90.
- Gaaf, W. van der. 1929. The conversion of the Indirect Personal Object into the Subject of a Passive Construction. *English Studies* 11: 1–10 and 58–67.
- Gaaf, W. van der. 1930. The passive of a verb accompanied by a preposition. *English Studies* 12: 1–24.
- Gelderen, E. van. 2011. Valency changes in the history of English. *Journal of Historical Linguistics* 1(1): 106–143.
- Givón, T. 1979. *On Understanding Grammar*. New York: Academic Press.
- Givón, T. 1983. Topic continuity in discourse: an introduction. In Givón, T. (ed.), *Topic continuity in discourse*, 4–41. Amsterdam: John Benjamins.
- Givón, T. 1984. *Syntax: A functional-typological introduction. Vol. 1*. Amsterdam: John Benjamins.
- Givón, T. 1990. *Syntax: A Functional-Typological Introduction. Vol. 2*. Amsterdam: John Benjamins.
- Givón, T. 2001. *Syntax: A Functional-Typological Introduction, Vol. 2* [Revised edition of Givón 1990]. Amsterdam: John Benjamins.
- Givón, T. 2006. Grammatical relations in passive clauses: A diachronic perspective. In Abraham, W. & Leisiö, L. (eds.), *Passivization and typology: Form and function*, 337–350. Amsterdam: John Benjamins.
- Godfrey, J., Holliman, E. & McDaniel, J. 1992. Switchboard: Telephone speech corpus for research and development. *International Conference on Acoustics, Speech and Signal Processing* 92, Vol. I, 517–520.
- Goh, G-Y. 2000. *The Synchrony and Diachrony of the English Prepositional Passive: Form, Meaning and Function*. PhD dissertation, Ohio State University
- Goh, G-Y. 2001. The Advent of the Prepositional Passive: An Innovation of Middle English? *English Studies* 82(3): 203–217.
- Gregory, M.L., & Michaelis, L.A. 2001. Topicalization and left-dislocation: A functional opposition revisited. *Journal of Pragmatics* 33: 1665–1706.
- Gundel, J.K. 1988. Universals of topic-comment structure In Hammond, M., Moravczik, E. & Wirth, J. (eds.), *Studies in Syntactic Typology*, 209–239. Amsterdam: John Benjamins.
- Gundel, J. K., N. Hedberg and R. Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69: 274–307.

- Haeberli, E. 2000. Adjuncts and the Syntax of Subjects in Old and Middle English. In Pintzuk, S., Tsoulas, G. & Warner, A. (eds.), *Diachronic Syntax: Models and Mechanisms*, 109–131. Oxford: Oxford University Press.
- Haeberli, E. 2002. Observations on the Loss of Verb Second in the History of English. In Zwart, C.J.W. & Abraham, W. (eds.), *Studies in Comparative Syntax*, 245–272. Amsterdam: John Benjamins.
- Haeberli, E. 2007. The development of subject-verb inversion in Middle English and the role of language contact. *GG@G (Generative Grammar in Geneva)* 5:15–33.
- Hale, K. & Keyser, S.J. 1993. On argument structure and the lexical expression of syntactic relations. In Hale, K. & Keyser, S.J. (eds.), *The View from Building 20*, 53–109. Cambridge, Mass.: MIT Press.
- Hale, K. & Keyser, S.J. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge, Mass.: MIT Press.
- Halliday, M.A.K. 1967. Notes on transitivity and theme in English: Part 2. *Journal of Linguistics* 3(2): 199–244.
- Halliday, M.A.K. 1968. Notes on transitivity and theme in English: Part 3. *Journal of Linguistics* 4(2): 179–215.
- Halliday, M.A.K. 1994. *An introduction to functional grammar*. London: Edward Arnold.
- Halliday, M.A.K. 2000. Grammar and Daily Life: Concurrence and Complementarity. In Lookwood, D.G., Fries, P. & Copeland, J. (eds.), *Functional Approaches to Language, Culture and Cognition: Papers in Honour of Sydney Lamb (CILT163)*, 221–237. Amsterdam: John Benjamins.
- Hallman, P. & Kallulli, D. 2013. Introduction to the Special Issue (Passive: Components and Composition). *Lingua* 125: 1–6.
- Hannay, M. & Keizer, E. 1993. Translation and contrastive grammar: The grammatical versus the communicative strategy. *Toegepaste Taalwetenschap in Artikelen* 45: 65–88.
- Haspelmath, M. 1992. Grammaticization of passive morphology. *Studies in Language* 14: 25–72.
- Haspelmath, M. 1993. More on the Typology of the Inchoative/Causative Verb Alternations. In Comrie, B. & Polinsky, M. (eds.), *Causatives and Transitivity*, 87–120. Amsterdam: John Benjamins.
- Hasselgård, H. 2010. *Adjunct Adverbials in English*. Cambridge: Cambridge University Press.
- Haug, D.T.T., Jøhndal, M.L., Eckhoff, H.M., Welo, E., Hertenberg, M.J.B & Muth, A. 2009. Computational and linguistic issues in designing a syntactically annotated parallel corpus of Indo-European languages. *TAL* 50(2): 17–45.
- Haug, D.T.T., Eckhoff, H.M. & Welo, E. 2014. The theoretical foundations of givenness annotation. In Bech, K. & Eide, K.G. (eds.), *Information Structure and Syntactic Change in Germanic and Romance Language*, 17–52. Amsterdam: John Benjamins.
- Haviland, S.E. & Clark, H.H. 1974. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning & Verbal Behavior* 13: 512–521.
- Hawkins, J. 1986. *A Comparative Typology of English and German*. London: Croom Helm.
- Hawkins, J. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hermodsson, L. 1952. *Reflexive und intransitive Verba im ältern Westgermanischen*. Uppsala: Almqvist & Wiksells.
- Hinds, J. 1975. Passives, pronouns and themes and rhemes. *Glossa* 9(1): 79–106.

- Hinnebusch, T.J. & Mirza, S.M. 1998. *Kiswahili: Msingi wa kusema, kusoma, na kuandika (Swahili: A Foundation for Speaking, Reading and Writing)*. Lanham, Md.: University Press of America.
- Hinterhölzl, R. & Petrova, S. 2010. From V1 to V2 in West Germanic. *Lingua* 120: 315–328.
- Hoekstra, T. & Mulder, R. 1990. Unergatives as Copular Verbs; Locational and Existential Predication. *The Linguistic Review* 7(1): 1–79.
- Hoekstra, T. & Roberts, I. 1993. Middle constructions in Dutch and English. In Reuland, E. & Abraham, W. (eds.), *Knowledge and language, Vol. 2: Lexical and conceptual structure*, 183–220. Dordrecht: Kluwer.
- Hoffmann, G. 1934. *Die Entwicklung des umschriebenen Perfektums im Altenglischen und Frühmittelenglischen*. PhD dissertation, University of Breslau.
- Hornstein, N. & Weinberg, A. 1981. Case theory and preposition stranding. *Linguistic Inquiry* 12(1): 55–91.
- Huddleston, R. D. & Pullum, G.K. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Huddleston, R.D. 1984. *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.
- Hundt, M. 2007. *English Mediopassive Constructions: A Cognitive, Corpus-Based Study of Their Origin, Spread and Current Status*. Amsterdam: Rodopi.
- Hutchins, W. J. 1975. Subjects, Themes, and Case Grammars. *Lingua* 35(2): 101–133.
- Inada, T. 1981. Problems of reanalysis and preposition stranding. *Studies in English Linguistics* 9: 121–130
- Jacobsson, B. 1951. *Inversion in English: with Special Reference to the Early Modern English Period*. Uppsala: Almqvist and Wiksell.
- Jaeggli, O. 1986. Passive. *Linguistic Inquiry* 17(4): 587–622
- Jespersen, O. 1909–49. *A modern English grammar on historical principles*. Copenhagen: Munksgaard.
- Jespersen, O. 1924. *The Philosophy of Grammar*. New York: Norton. Reprinted in 1965.
- Jespersen, O. 1927. *A modern English grammar on historical principles. Part III: syntax, vol. 2*. Heidelberg: Winter.
- Jones, C. 1987. P for proper governor. *West Coast Conference on Formal Linguistics* 6, 115–130.
- Karttunen, L. 1976. Discourse referents. In McCawley, J.D. (ed.), *Syntax and semantics 7: Notes from the linguistic underground*, 363–385. New York: Academic Press.
- Kemenade, A.M.C. van. 1987. *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.
- Kemenade, A.M.C. van. 1997. V2 and Embedded Topicalisation in Old and Middle English. In Kemenade, A.M.C. van & Vincent, N. (eds.). *Parameters of Morphosyntactic Change*, 326–352. Cambridge: Cambridge University Press.
- Kemenade, A.M.C. van. 1998. Review of Cynthia Allen, *Case Marking and Reanalysis*, Oxford University Press 1995. *Journal of Linguistics* 34: 227–232.
- Kemenade, A.M.C. van. 2012. Rethinking the loss of verb second. In: Nevalainen, T. & Closs-Traugott, E. (eds.), *The Oxford Handbook of the History of English*, 822–834. Oxford: Oxford University Press.
- Kemenade, A.M.C. van & Los, B.L.J. 2006. Discourse Adverbs and Clausal Syntax in Old and Middle English. In Kemenade, A.M.C. van & Los, B.L.J. (eds.), *The Handbook of the History of English*, 224–248. Malden: Blackwell.

- Kemenade, A.M.C. van, Milićev, T. & Baayen, H. 2008. The balance between discourse and syntax in Old and Middle English. In Dossena, M., Gotti, M. & Dury, R. (eds.), *English Historical Linguistics 2006: Selected papers from the fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21–25 August 2006, Vol 1: Syntax and Morphology*, 3–22. Amsterdam: John Benjamins.
- Kemenade, A.M.C. van & T. Milićev. 2012. Syntax and Discourse in Old English and Middle English Word Order. In Donas, D., Whitman, J. & Garrett, A. *Grammatical Change: Origins, Nature, Outcomes*, 239–255. Oxford: Oxford University Press.
- Kemenade, A.M.C. van & Westergaard, M. 2012. Syntax and Information Structure: Verb Second variation in Middle English. In Meurman-Solin, A., Lopez-Couso, M. J. & Los, B.L.J. (eds.), *Information Structure and Syntactic Change in the History of English* (Oxford Studies in the History of English 1), 87–118. Oxford: Oxford University Press.
- Kenyon, J. 1909. *The syntax of the infinitive in Chaucer* (Chaucer Society Publications, 2nd series, no. 44). London: K. Paul, Trench, Trübner and Co.
- Keyser, S.J. & Roeper, T. 1984. On the middle and ergative constructions in English. *Linguistic Inquiry* 15: 381–416.
- Kilpiö, M. 1989. *Passive constructions in Old English translations from Latin: With special reference to the OE Bede and the Pastoral Care* (Mémoires de la Société Néophilologique de Helsinki 49). Helsinki: Société néophilologique.
- Kirkwood, H.W. 1978. Options and constraints in the surface ordering of noun phrases in English and German. *Journal of Pragmatics* 2: 225–245.
- Kisbye, T. 1972. *An Historical Outline of English Syntax. I-II*. Aarhus: Akademisk Boghandel.
- Klaeber, Friedrich. 1923. Eine Bemerkung zum altenglischen Passivum. *Englische Studien* 57: 187–95.
- Klingebiel, J. 1937. *Die Passivumschreibungen im Altenglischen*. Böttrop: Postberg.
- Kohonen, T. 2007. From Helsinki through the centuries: The design and development of English diachronic corpora. In Pahta, P., Taavitsainen, I., Nevalainen, T. & Tyrkkö, J. (eds.), *Towards Multimedia in Corpus Studies (Studies in Variation, Contacts and Change in English 2)*. (<http://www.helsinki.fi/varieng/series/volumes/02/kohonen/>)
- Komen, E.R. 2011. *CorpusStudio*. Nijmegen: Radboud University.
- Komen, E.R., Hebing, R.G.A., Kemenade, A.M.C. van & Los, B.L.J. 2014. Quantifying information structure changes in English. In Bech, K. & Eide, K.G. (eds.), *Information Structure and Syntactic Change in Germanic and Romance Languages*, 81–110. Amsterdam: John Benjamins.
- Koopman, W. 1985. The syntax of verb and particle combinations in Old English. In Bennis, H. & Beukema, F. (eds.), *Linguistics in the Netherlands 1985*, 91–100. Dordrecht: Foris.
- Kratzer, A., 1996. Severing the external argument from its verb. In Rooryck, J. & Zaring, L. (eds.), *Phrase Structure, the Lexicon*, 109–137. Dordrecht: Springer.
- Krifka, M. 2003. Semantic and pragmatic conditions for the Dative Alternation. *Korean Journal of English Language and Linguistics* 4: 1–32.
- Krifka, M. 2007. Basic notions of information structure. In Féry, C., Fanselow, G. & Krifka, M. (eds.), *The Notions of Information Structure*, 13–55. Potsdam: Potsdam Universitätsverlag.
- Kroch, A. & Taylor, A. 1997. Verb movement in Old and Middle English: Dialect variation and language contact. In Kemenade, A.M.C. van & Vincent, N. (eds.), *Parameters of morphosyntactic change*, 297–325. Cambridge: Cambridge University Press.
- Kroch, A. & Taylor, A. 2000. *Penn-Helsinki Parsed Corpus of Middle English, second edition*. (<http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-3/index.html>)

- Kroch, A. 1989. Reflexes of grammar in patterns of language change. *Language variation and change* 1: 199–244.
- Kroch, A., Santorini, B. & Delfs, L. 2004. *The Penn-Helsinki Parsed Corpus of Early Modern English* (PPCEME). Department of Linguistics, University of Pennsylvania. (<http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-2/index.html>)
- Kroch, A., Santorini, B. & Dierani, A. 2010. *The Penn-Helsinki Parsed Corpus of Modern British English* (PPCMBE). Department of Linguistics, University of Pennsylvania. (<http://www.ling.upenn.edu/hist-corpora/PPCMBE-RELEASE-1/index.html>)
- Kruijff-Korbayová, I. & Steedman, M. 2003. Discourse and Information Structure. *Journal of Logic, Language and Information* 12: 249–259.
- Kucera, H., & Francis, W.N. 1967. *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Kuno, S. & Takami, K. 2004. Functional Constraints in Grammar: On the unergative-unaccusative distinction. Amsterdam: John Benjamins.
- Kuno, S. 1979. On the Interaction between Syntactic Rules and Discourse Principles. Unpublished manuscript.
- Kytö, M. & Pahta, P. 2012. Evidence from historical corpora up to the twentieth century. In Nevalainen, T. & Traugott, E. C. (eds.). *The Oxford Handbook of the History of English* (Oxford Handbooks in Linguistics), 123–133. New York: Oxford University Press.
- Lambrecht, K. 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Larson, R.K. 1988. On the Double Object Construction. *Linguistic Inquiry* 19(3): 335–391.
- Lee, P-H. 2007. A Morphological Account of the Emergence of Prepositional Passives. *Studies in Modern English Grammar* 49: 99–126.
- Levin, B. and Rappaport Hovav, M. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface* (Linguistic Inquiry Monograph 26). Cambridge, Mass.: MIT Press.
- Li, C.N. & Thompson, S.A. 1976. Subject and topic: a new typology of language. In Li, C.N. (ed.), *Subject and Topic*, 457–461. Austin: University of Texas Press.
- Lieber, R. 1979. The English Passive: An Argument for Historical Rule Stability. *Linguistic Inquiry* 10: 667–688.
- Lightfoot, D.W. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.
- Lightfoot, D.W. 1981. The history of NP movement. In Baker, C.L. & McCarthy, J. (eds.), *The logical problem of language acquisition*, 86–119. Cambridge, Mass.: MIT Press.
- Los, B.L.J. 1999. Infinitival Complementation in Old and Middle English (LOT Dissertation Series 31). PhD dissertation, Vrije Universiteit Amsterdam. The Hague: Thesus.
- Los, B.L.J. 2002. The loss of the indefinite pronoun man: Syntactic change and information structure. In Fanego, T., López-Couso, M.J. & Pérez-Guerra, J. (eds.), *English Historical Syntax and Morphology* (Current Issues in Linguistic Theory Series 223), 181–202. Amsterdam: Benjamins.
- Los, B.L.J. 2004. From resultative predicate to event-modifier. In Kay, C., Horobin, S. & Smith, J.J. (eds.), *New Perspectives on English Historical Linguistics: Selected papers from 12 ICEHL, Glasgow, 21–26 August 2002*, 83–102. Volume I: Syntax and Morphology.
- Los, B.L.J. 2005. *The rise of the to-infinitive*. Oxford: Oxford University Press.
- Los, B.L.J. 2009. The consequences of the loss of verb-second in English: information structure and syntax in interaction. *English Language and Linguistics* 13(1): 97–125.
- Los, B.L.J. 2012. The Loss of Verb-Second and the Switch from Bounded to Unbounded Systems. In Meurman-Solin, A., Lopez-Couso, M. J. & Los, B.L.J. (eds.), *Information*

- Structure and Syntactic Change In the History of English*, 21–46. Oxford: Oxford University Press.
- Los, B.L.J. 2013. Recycling "Junk": A case for exaptation as a response to breakdown. In Kikusawa, R. & Reid, L. A. (eds.), *Historical Linguistics 2011: Selected papers from the 20th International Conference on Historical Linguistics*, Osaka, 25-30 July 2011 (Current Issues in Linguistic Theory 326), 267–288. Amsterdam: John Benjamins.
- Los, B.L.J. & Dreschler, G.A. 2012. The loss of local anchoring: From adverbial local anchors to permissive subjects. In Nevalainen, T. & Traugott, E. C. (eds.). *The Oxford Handbook of the History of English* (Oxford Handbooks in Linguistics), 859–871. New York: Oxford University Press.
- Los, B.L.J. & Komen, E.R. 2012. Clefts as resolution strategies after the loss of a multifunctional first position. In Nevalainen, T. & Traugott, E. C. (eds.). *The Oxford Handbook of the History of English* (Oxford Handbooks in Linguistics), 884–898. New York: Oxford University Press.
- Los, B.L.J., Blom, C., Booij, G., Elenbaas, M. & Kemenade, A.M.C. 2012. *Morphosyntactic Change: A comparative study of particles and prefixes* (Cambridge Studies in Linguistics; no. 134). Cambridge: Cambridge University Press.
- Los, B.L.J., Verheijen, L., & Haan, P. de. 2013. Information Structure: The final hurdle?: The development of syntactic structures in (very) advanced Dutch EFL writing. *Dutch Journal of Applied Linguistics* 2: 92–107.
- Mair, C. 1990. *Infinitival complement clauses in English: A study of syntax in discourse*. Cambridge: Cambridge University Press
- Marantz, A., 1984. *On the Nature of Grammatical Relations*. Cambridge, Mass.: MIT Press.
- Marcus, M., Santorini, B. & Marcinkiewicz, M.A. 1993. Building A Large Annotated Corpus of English: The Penn Treebank. *Technical Reports (CIS). Paper 237*.
- Mathesius, V. 1928. On linguistic characterology with illustrations from Modern English. In *Actes du Premier Congrès International de Linguistes*, The Hague. Reprinted in Vachek, J. (ed.). 1964. *A Prague School Reader in Linguistics*, 59–74. Bloomington IN: Indiana University Press.
- Mathesius, V. 1929. Functional linguistics. In Vachek, J.(ed.), *Praguiana: Some Basic and Less Well-Known Aspects of the Prague Linguistics School*, 121–142. Amsterdam: John Benjamins.
- McColl Millar, R. 2000. Some suggestions for explaining the origin and development of the definite article in English. In Fischer, O.C.M., Rosenbach, A. & Stein, D. (eds.), *Pathways of Change: Grammaticalization in English*, 275–310. Amsterdam: John Benjamins.
- McFadden, T. & Alexiadou, A. 2006. Pieces of the perfect in German and older English. In Baumer, D. Montero, D. & Scanlon, M. (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 270–278. Somerville, Mass.: Cascadilla Proceedings Project.
- McMillion, A. 2006. *Labile Verbs in English*. PhD dissertation, Stockholm University.
- Mitchell, Bruce. 1985. *Old English Syntax, Vol. 1: Concord, the Parts of Speech, and the Sentence*. Oxford: Clarendon.
- Moessner, L. 1994. Early Modern English passive constructions. In Kastovsky, D. (ed.), *Studies in Early Modern English*, 217–232. Berlin: Mouton de Gruyter.
- Mustanoja, T.F. 1960. *A Middle English syntax; Part 1: parts of speech*. Helsinki: Société Néophilologique.
- Newmeyer, F.J. 2007. 'More complicated and hence, rarer': A look at grammatical complexity and cross-linguistic rarity. In Karimi, S., Samiian, V. & Wilkins, W.K. (eds.), *Phrasal and*

- Clausal Architecture: Syntactic derivation and interpretation. In honor of Joseph E. Emonds, 221–242. Amsterdam: John Benjamins.
- Noël, D. & Auwera, J. van der. 2009. Revisiting <be supposed to> from a diachronic constructionist perspective. *English Studies* 90(5): 599–623.
- Noël, D. 1998. Infinitival copular complement clauses in English: Explaining the predominance of passive matrix clauses. *Linguistics* 36(6): 1045–1063.
- Noël, D. 2001. The passive matrices of English infinitival complement clauses: Evidentials on the road to auxiliarihood? *Studies in Language* 25(2): 255–296.
- Noël, D. 2003. The be said to construction in Late Modern English. In Remael, A. & Pelsmaekers, K. (eds.), *Configurations of culture*, 155–164. Antwerpen: Garant.
- Noël, D. 2008. The nominative and infinitive in Late Modern English: A diachronic constructionist approach. *Journal of English Linguistics* 36(4): 314–340.
- Oehrle, R. 1979. *The Grammatical Status of the English Dative Alternation*. PhD dissertation, MIT, Cambridge, Mass.
- Ohkado, M. 2004. Coordinate clauses in Old English with special reference to Ælfric's Catholic Homilies. *Folia Linguistica Historica* 25(1-2): 155–176.
- Peitsara, K. 1992. On the development of the *by*-agent in English. In Rissanen, M., Ihalainen, O., Nevalainen, T. & Taavitsainen, I. (eds.), *History of Englishes: new methods and interpretations in historical linguistics* (Topics in English Linguistics 10), 379–400. Berlin: Mouton de Gruyter.
- Perlmutter, D. M. 1978. Impersonal passives and the Unaccusative Hypothesis. In *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, 157–189.
- Pesetsky, D. & Torrego, E. 2004. The Syntax of Valuation and the Interpretability of Features. In Karimi, S., Samiian, V. & Wilkins, W.K. (eds.), *Phrasal and Clausal Architecture: Syntactic derivation and interpretation. In honor of Joseph E. Emonds*, 262–294. Amsterdam: John Benjamins.
- Petré, P. & Cuyckens, H. 2008. The Old English copula *weorðan* and its replacement in Middle English. In Gotti, M. Dossena, M. & Dury, R. (eds.), *English Historical Linguistics 2006: Volume I: Syntax and Morphology*, 23–48. Amsterdam: John Benjamins.
- Petré, P. 2010. *On the interaction between constructional & lexical change: Copular, Passive and related Constructions in Old and Middle English*. PhD dissertation, Catholic University Leuven.
- Pintzuk, S. 1991. *Phrase structures in competition: Variation and change in Old English word order*. PhD dissertation, University of Pennsylvania.
- Pintzuk, S. 1993. Verb seconding in Old English: verb movement to Infl. *The linguistic review* 10: 5–35.
- Platzack, C. 2005. Cross-Germanic Promotion to Subject in Ditransitive Passives: a Feature-Driven Account. In Vulchanova, M. & Åfarli, T. A. (eds.), *Grammar and Beyond: essays in honour of Lars Hellan*, 135–161. Oslo: Novus Press.
- Pollock, J.Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.
- Postal, P.M. 1974. *On Raising: One rule of English grammar and its theoretical implications*. Cambridge, Mass.: MIT Press.
- Prince, E.L. 1981. Towards a taxonomy of Given-New Information. In Cole, P. (ed.), *Radical Pragmatics*, 223–56. New York: Academic Press.
- Prince, E.L. 1992. The Zpg Letter: Subjects, Definiteness, and Information-Status. In Thompson, S. & Mann, W. (eds.), *Discourse Description: Diverse Analyses of a Fund-Raising Text*, 195–325. Amsterdam: John Benjamins.

- Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. 1972. *A Grammar of Contemporary English*. London: Longman.
- Randall, Beth. 2005-2007. *CorpusSearch 2*. (<http://corpussearch.sourceforge.net>)
- Rappaport Hovav, M. & Levin, B. 2008. The English Dative Alternation: The Case for Verb Sensitivity. *Journal of Linguistics* 44: 129–167.
- Richards, M. 2004. *Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax*. PhD dissertation, University of Cambridge.
- Riemsdijk, H.C. van. 1978. *A case study in syntactic markedness: the binding nature of prepositional phrases* (Studies in Generative Grammar 4). Lisse: Peter de Ridder.
- Rissanen, M. 1994. The position of *not* in Early Modern English questions. In Kastovsky, D. (ed.), *Studies in Early Modern English*, 339–48. Berlin: Mouton de Gruyter.
- Rissanen, M. 1999. Isn't it? or is it not? On the order of postverbal subject and negative particle in the history of English. In Tiekens-Boon van Ostade, I., Tottie, G. & Van der Wurff, W. (eds.), *Negation in the History of English*, 189–205. Berlin: Mouton de Gruyter.
- Rizzi, L. 1997. The fine structure of the left periphery. In Haegeman, L. (ed.), *Elements of Grammar: Handbook in Generative Syntax*. Dordrecht: Kluwer.
- Rochemont, M. 1978. *A Theory of Stylistic Rules in English*. PhD dissertation, University of Massachusetts, Amherst.
- Rohdenburg, G. 1974. Sekundäre Subjektivierungen im Englischen und Deutschen: Vergleichende Untersuchungen zur Verb- und Adjektivsyntax. *PAKS-Arbeitsbericht* 8. Bielefeld: Cornelson-Velhagen and Klasing.
- Russell, B. 1905. On denoting. *Mind* 14: 479–493.
- Russom, J.H. 1982. An Examination of the Evidence for OE Indirect Passives. *Linguistic Inquiry* 13(4): 677–680.
- Sato, K. 2012. Ælric's linguistic and stylistic alterations in his adaptations from the Old English *Boethius*. *Neophilologus* 96(4): 631–640.
- Savage, A. & Watson, N. 1991. *Anchoritic spirituality: Ancrene Wisse and associated works*. New York: Paulist Press.
- Schippers, R. *The acquisition of particle verbs in Dutch*. PhD dissertation, Utrecht University. Utrecht: LOT.
- Schmidt, D. 1980. *A History of Inversion in English*. Dissertation, Ohio State University.
- Seoane, E. 1999a. Inherent topicality and object foregrounding in Early Modern English. *ICAME Journal* 23, 117–140.
- Seoane, E. 1999b. The consolidation of the indirect and prepositional passive in Early Modern English: Evidence from the Helsinki Corpus. *Estudios Ingleses de la Universidad Complutense* 7: 119–139.
- Seoane, E. 2000. The passive as an information-rearranging device in Early Modern English. *Studia Neophilologica* 72: 24–33.
- Seoane, E. 2006. Information structure and word order: The passive as an information rearranging strategy. In Van Kemenade, A.M.C. & Los, B.L.J. (eds.), *Handbook of the History of English*, 360–391. Oxford: Blackwell.
- Sgall, P., Hajicová, E. & Panevová, J. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Dordrecht: Reidel.
- Siewierska, A. & Hollmann, W.B. 2007. Ditransitive clauses in English with special reference to Lancashire dialect. In Hannay, M. & Steen, G.J. (eds.), *Structural-functional studies in English grammar: in honour of Lachlan Mackenzie (Studies in Language Companion Series)*, 83–102. Amsterdam: John Benjamins.
- Siewierska, A. 1984. *The Passive: A Comparative Linguistic Analysis*. London: Croom Helm.

- Siewierska, A. 2013. Passive Constructions. In Dryer, M.S. & Haspelmath, M. (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://wals.info/chapter/107>) (Accessed 2 July 2014)
- Skeat, W.W. 1966 [1881-1900]. *Ælfric's Lives of Saints* (Early English Text Society 76, 82, 94, 114). London: Oxford University Press.
- Speyer, A. 2010. *Topicalization and Stress Clash Avoidance in the History of English* (Topics in English Linguistics 69). Berlin: De Gruyter Mouton.
- Stokoe, H. R. 1937. *The understanding of syntax*. London: Heinemann.
- Stroik, T. 1992. Middles and movement. *Linguistic Inquiry* 23: 127–137.
- Svartvik, J. 1966. *On Voice in the English Verb*. The Hague: Mouton.
- Sweet, H. 1898. *A New English Grammar, Logical and Historical; Part II, Syntax*. London: Clarendon Press.
- Sweet, H. 1900. *A New English Grammar, Logical and Historical; Part I*. London: Clarendon Press.
- Takami, K-I. 1992. *Preposition Stranding: From Syntactic to Functional Analyses*. Berlin, New York: de Gruyter.
- Tatlock, J. & Kennedy, A. 1927. *A concordance to the Complete Works of Geoffrey Chaucer and to the Romaunt of the Rose*. Washington: The Carnegie Institution of Washington.
- Taylor, A. & Pintzuk, S. 2011. The interaction of syntactic change and information status effects in the change from OV to VO in English. *Catalan Journal of Linguistics* 10: 71–94.
- Taylor, A. & Pintzuk, S. 2012a. The Effect of Information Structure on Object Position in Old English: a Pilot Study. In Lopez-Couso, M-J., Los, B.L.J. & Meurman-Solin, A. (eds.), *Information structure and syntactic change in the history of English*, 47–65. Oxford: Oxford University Press.
- Taylor, A. & Pintzuk, S. 2012b. Rethinking the OV/VO alternation in Old English: The effect of complexity, grammatical weight, and information status. In Nevalainen, T. & Traugott, E. (eds.), *The Oxford Handbook of the History of English* (Oxford Handbooks in Linguistics), 835–845. Oxford: Oxford University Press.
- Taylor, A. & Pintzuk, S. 2014. Testing the theory: Information structure in Old English. In Bech, K. & Eide, K.G. (eds.), *Information Structure and Syntactic Change in Germanic and Romance Language*, 53–77. Amsterdam: John Benjamins.
- Taylor, A., Warner, A.R., Pintzuk, S. & Beths, F. 2003. *The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE)*. Department of Linguistics, University of York. (<http://www-users.york.ac.uk/~lang22/YcoeHome1.htm>)
- Thompson, S.A. 1987. The Passive in English: A Discourse Perspective. In Channon, R. & Shockey, L. (eds.), *In Honor of Ilse Lehiste: Ilse Lehiste Puhendusteos*, 497–511. Dordrecht: Foris
- Thornburg, L. 1985. The history of the prepositional passive in English. *Proceedings of the annual meeting of the Berkeley Linguistics Society* 11: 327–336.
- Thorpe, I. 1846. *The Homilies of the Anglo-Saxon Church: The First Part, Containing the Sermones Catholici*, Vol. 2. Reprinted in 2013, Cambridge: Cambridge University Press.
- Toole, J. 1996. The effect of genre on referential choice. In Fretheim, T. & Gundel, J.K. (eds.), *Reference and Referent Accesibility*, 263–90. Amsterdam: Benjamins.
- Toyota, J. 2008. *Diachronic change in the English passive*. Basingstoke: Palgrave Macmillan.
- Vallduví, E. 1990. *The informational component*. PhD dissertation, University of Pennsylvania.
- Vikner, S. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: Oxford University Press.

- Virtanen, T. 1992. *Discourse Functions of Adverbial Placement in English: Clause-Initial Adverbials of Time and Place in Narratives and Procedural Place Descriptions*. Abo, Finland: Abo Akademi University Press.
- Visser, F.T. 1963-73. *An historical syntax of the English language*. 3 parts; 4 vols. Leiden: E.J. Brill.
- Visser, F.T. 1973. *An historical syntax of the English language, Part III.2*. Leiden: E.J. Brill.
- Walkden, G. 2015. Verb-third in early West Germanic: a comparative perspective. In Biberauer, T. & Walkden, G. (eds.), *Syntax over time: lexical, morphological and information-structural interactions*, 236–248. Oxford: Oxford University Press.
- Wanner, A. 2009. *Deconstructing the English Passive* (Topics in English Linguistics 41). Berlin: De Gruyter Mouton.
- Warner, A.R. 1982. *Complementation in Middle English and the Methodology of Historical Syntax*. London: Croom Helm.
- Warner, A.R. 2007. Parameters of variation between verb-subject and subject-verb order in late Middle English. *English Language and Linguistics* 11: 81–111.
- Wasow, T. 1977. Transformations and the lexicon. In Culicover, P.W., Wasow, T. & Akmajian, A. (eds.), *Formal syntax*, 327–60. New York: Academic Press.
- Wasow, T. 2002. *Postverbal behavior*. Stanford: CSLI Publications.
- Wattie, J.M. 1930. Tense. *English Studies* 16: 121–143.
- Zeitlin, J. 1908. *The ACI and Some Kindred Constructions in English*. PhD dissertation, University of Colombia.
- Zwart, C.J.W. 1997. *Morphosyntax of verb movement: a Minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer Academic Publishers.

Samenvatting (Summary in Dutch)

In dit proefschrift onderzoek ik de relatie tussen twee onderdelen van het taalsysteem, syntaxis en informatiestructuur, door het bestuderen van twee opvallende veranderingen in de geschiedenis van het Engels. De eerste verandering is het verval van de zogenoemde verb-secondregel (V2) in de veertiende en vijftiende eeuw, en de tweede verandering is de introductie van nieuwe typen lijdende of passieve vormen in dezelfde periode. Hierbij wordt syntaxis gezien als het onderdeel dat bepaalt welke zinsconstructies mogelijk zijn in een taal en informatiestructuur als het onderdeel dat de optimale volgorde van informatie in een zin bepaalt. Bij het verval van V2 en de introductie van nieuwe types passieve vormen spelen zowel de syntaxis als de informatiestructuur een belangrijke rol. De vraag die ik probeer te beantwoorden is of deze twee onderdelen van het taalsysteem elkaar beïnvloeden of onafhankelijk van elkaar opereren. De studies in dit proefschrift laten zien dat de syntaxis de informatiestructuur beïnvloedt, want het is de syntaxis die bepaalt welke grammaticale opties beschikbaar zijn in de taal. Voor een omgekeerde invloed is weinig bewijs—het is niet zo dat de informatiestructuur de oorzaak is van het ontstaan van nieuwe constructies—maar de informatiestructuur is wel heel bepalend voor hoe de beschikbare constructies gebruikt worden. Dit geldt zowel voor de frequentie van verschillende zinstypen als voor de precieze manier waarop deze zinstypen worden gebruikt.

Twee grote veranderingen in de veertiende en vijftiende eeuw

In de veertiende en vijftiende eeuw maakte het Engels twee opmerkelijke veranderingen door: het verval van de zogenoemde V2-regel en de introductie van nieuwe passieve vormen. De eerste van deze twee veranderingen, het verval van de V2-regel, gaat om de woordvolgorde in hoofdzinnen. De V2-regel bepaalt dat de persoonsvorm altijd op de tweede positie staat in de hoofdzin, zoals dat in het hedendaags Nederlands en Duits ook is.

- (1) Ik ging gisteren naar Londen.
- (2) Gisteren ging ik naar Londen.

In voorbeeld (1) begint de zin met een onderwerp, dat gevolgd wordt door de persoonsvorm (het vervoegde werkwoord). In voorbeeld (2) begint de zin met een tijdsbepaling, maar de persoonsvorm is nog steeds het tweede element in de zin, terwijl het onderwerp de persoonsvorm hier volgt. Deze omgekeerde volgorde van onderwerp en persoonsvorm wordt inversie genoemd. Het verschil met het hedendaags Engels is dat als er in het hedendaags Engels een ander element dan het onderwerp op de eerste positie in de Engelse zin staat, de volgorde van onderwerp en persoonsvorm niet verandert, zoals voorbeeld (3) laat zien.

- (3) Last week I went to Manchester.
Vorige week ik ging naar Manchester
'Vorige week ging ik naar Manchester'

In (3) staat het onderwerp op de tweede positie, voor de persoonsvorm. Maar hoewel in het modern Engels het onderwerp praktisch altijd vóór de persoonsvorm staat, was er in het Oudengels (450-1150) een variant van de V2-regel in gebruik (eerst aangetoond door Van Kemenade 1987). Een belangrijk verschil met het Nederlands en Duits is dat het Oudengels een niet volledig regelmatig systeem heeft, maar dat er—afhankelijk van het element in eerste positie—variatie is. Na de bijwoorden *þa* en *þonne* is de volgorde bijvoorbeeld altijd *persoonsvorm-onderwerp*, maar na andere tijdsbepalingen en lijdend voorwerpen komen zowel *persoonsvorm-onderwerp* als *onderwerp-persoonsvorm* voor. Uit eerder onderzoek naar V2 is gebleken dat deze woordvolgorde vanaf ongeveer 1400 begint af te nemen, en bijna geheel verdwenen is rond 1500. Uit deze studies blijkt overigens ook dat er niet alleen veel variatie is in het systeem zoals dat in het Oudengels gebruikt wordt, maar ook in het verval van V2: in verschillende contexten neemt inversie in verschillende snelheden af.

Ook in het gebruik en de historische ontwikkeling van de passieve vorm is het Engels bijzonder in vergelijking met nauwverwante talen zoals Nederlands en Duits. Verschillende taalkundigen hebben beweerd dat in het hedendaags Engels de passieve vorm vaker wordt gebruikt dan in deze verwante talen. Daarnaast laat een studie van Seoane (1999a) zien dat dit een gevolg lijkt te zijn van een toename in

gebruik van passieve vormen in de vijftiende en zestiende eeuw. Een nog meer opvallende ontwikkeling is dat het Engels meer passieve vormen toelaat dan deze verwante talen. Het Engels heeft, net als het Nederlands en Duits, een standaard passieve vorm waarbij het lijdend voorwerp van een actieve (of bedrijvende) zin het onderwerp wordt van de passieve zin, met de passieve markerings op het werkwoord.

- (4) a. The man bites the dog.
b. De man bijt de hond.
- (5) a. The dog is bitten by the man.
b. De hond wordt gebeten door de man.

Maar naast deze standaardvorm heeft het Engels ook de volgende vormen.

- (6) a. Someone sat on this bench.
'Iemand zat op dit bankje'
b. This bench was sat on.
dit bankje was gezeten op
'Er is op dit bankje gezeten'
- (7) a. John gave her the book
John gaf haar het boek
b. She was given the book
zij werd gegeven het boek
'Aan haar werd het boek gegeven'
- (8) a. They believed him to be a liar.
zij geloofden hem te zijn een leugenaar
'Zij dachten dat hij een leugenaar was'
b. He was believed to be a liar.
hij werd geloofd te zijn een leugenaar
'Er werd van hem gedacht dat hij een leugenaar was'

In de zogenoemde *prepositional passive* ('prepositionele passieve vorm') in (6)b wordt om het zelfstandig naamwoord dat volgt op een voorzetsel (*on this bench*) het onderwerp van de passieve zin. Voorbeeld (7)b laat een zogenaamde *recipient passive* ('ontvanger passieve vorm') zien, waarin het meewerkend voorwerp (vaak een 'ontvanger') het onderwerp is geworden van de passieve zin. Deze twee passieve vormen vallen op omdat in de meeste talen met passieve vormen het lijdend voorwerp het onderwerp van de passieve zin wordt. Het laatste voorbeeld (8) laat een zinstype zien dat ook heel beperkt in het Nederlands wordt gebruikt

(bijvoorbeeld met *geacht te zijn*). In deze zin is het onderwerp van de passieve zin weliswaar het lijdend voorwerp in de hoofdzin, maar het functioneert ook als onderwerp ('uitvoerder') van het werkwoord *to be*. Omdat de actieve zin in de literatuur een ECM constructie⁹⁷ wordt genoemd, heet deze passieve vorm de *ECM passive*. Uit eerder onderzoek is gebleken dat deze nieuwe passieve vormen voor het eerst aan het einde van de Middelenengelse periode (1150-1500) voorkomen in de manuscripten, verrassend genoeg ongeveer rond dezelfde tijd. De *prepositional passive* wordt vanaf de dertiende eeuw gebruikt, de *recipient passive* en de *ECM passive* vanaf halverwege de veertiende eeuw.

De relatie tussen het verval van de V2-regel en de veranderingen in de passieve vorm

Beide veranderingen—het verval van de V2-regel en de introductie van nieuwe passieve vormen—zijn van een type dat niet veel voorkomt en zijn dus op zichzelf al het bestuderen waard. Maar er is een belangrijke reden waarom het aannemelijk is dat deze twee ontwikkelingen aan elkaar gerelateerd zijn, namelijk de interactie tussen twee onderdelen van het taalsysteem: de syntaxis, of grammatica, en de informatiestructuur. Deze twee onderdelen spelen een rol voor een spreker als deze een efficiënte boodschap wil overbrengen: de zin moet grammaticaal correct zijn maar moet ook op een optimale manier de informatie overbrengen.

Er is al veel onderzoek gedaan naar de relatie tussen informatiestructuur en syntaxis. Uit deze onderzoeken is een aantal belangrijke functies of principes naar voren gekomen. Allereerst is er het *given-before-new* principe, die bepaalt dat een zin, of boodschap, begint met de kennis die als gegeven ('given') wordt geacht, en dat daarna pas de nieuwe ('new') informatie wordt toegevoegd. Bij dit principe gaat het zowel om de relatieve volgorde van de informatie ('given' voor 'new') als om de absolute informatie in de eerste positie van de zin. Deze positie wordt in de literatuur (beginnend bij Halliday 1967, 1994) wel de 'theme' genoemd, het startpunt van de boodschap. Hoewel deze positie gebruikt kan worden voor nieuwe informatie of voor een tegenstelling, is een belangrijke functie van deze positie het weergeven van zogenaamde *unmarked themes*. Dit zijn elementen die niet contrastief zijn en 'given' informatie bevatten, en zo aansluiten op de vorige zin. Hierdoor ontstaat een samenhang in de tekst.

De positie vóór de persoonsvorm in het Oudengelse V2-systeem kan gebruikt worden door verschillende syntactische elementen, zoals het lijdend voorwerp, een bijwoord, of een voorzetselgroep (een voorzetsel met een zelfstandig naamwoord, bijvoorbeeld *op dit bankje*). Deze elementen kunnen verschillende types informatie bevatten, zowel bekende of 'gegeven' informatie als nieuwe informatie. Er is door onderzoekers ook aangetoond dat een deel van de variatie in de positie van het onderwerp bepaald wordt door de informatiestatus (of 'waarde') van dit onderwerp.

⁹⁷ ECM staat voor *Exceptional Case Marking*, een term van Chomsky (1981). Het lijdend voorwerp van de hoofdzin is tegelijkertijd het onderwerp van de bijzin en de naamval op dit lijdend voorwerp volgt niet de standaardregels.

Dit betekent dat het Oudengelse V2-systeem sprekers de mogelijkheid geeft om informatie te ordenen volgens het *given-before-new* principe en om gegeven informatie in de eerste positie in de zin te plaatsen. Deze eerste positie is belangrijk omdat dit de voorkeurspositie is voor ‘given’ informatie.

De lijdende vorm kan ook voor deze twee functies gebruikt worden. De passieve vorm verandert de volgorde van het onderwerp en de lijdende vorm zoals deze in de actieve zin staan (*Man bijt hond* wordt *Hond wordt gebeten door man*). Daarnaast zorgt een passieve vorm ervoor dat een argument dat niet oorspronkelijk een onderwerp is (het lijdend voorwerp) toch als onderwerp gebruikt kan worden.

Deze twee constructies—V2 en de passieve vorm—delen dus twee functies als het gaat om de informatiestructuur en daarom is het aannemelijk om te stellen dat als een van de constructies verdwijnt, dit gevolgen heeft voor de andere constructie. Dit is precies wat de tot nu toe bekende gegevens al suggereerden: V2 vervalt en de passief neemt toe, zowel in de frequentie waarin het wordt gebruikt als in de vormen die worden gebruikt. Deze relatie tussen passieve zinnen en het V2-systeem is eerder gelegd door Los (2005, 2009) en Seoane (2000, 2006). Los beargumenteert dat het verval van V2 niet alleen gaat om grammaticale veranderingen, maar eigenlijk een veel grotere verandering inhoudt voor de taal. Ze schrijft de introductie van de nieuwe passieve vormen toe aan het verval van V2: het verval van V2 zorgde voor minder flexibiliteit in de eerste positie (minder verschillende types elementen in eerste positie) en een toename van zinnen die beginnen met een onderwerp. Daarbij is dus vooral van belang dat de passieve vorm van een lijdend voorwerp een onderwerp kan maken. Seoane schrijft juist de toename in het gebruik van passieve vormen toe aan het verval van V2, niet alleen omdat het van lijdend voorwerpen een onderwerp maakt, maar ook omdat het de informatie in de zin kan herschikken volgens een *given-before-new* patroon. De vraag is dus of het verval van V2 inderdaad leidt tot nieuwe passieve vormen of dat op een wat indirectere manier de syntactische beperkingen die het verval van V2 inhoudt leiden tot een ander gebruik van overgebleven strategieën om toch in de informatiestructuurbehoeften van sprekers te kunnen voorzien.

Introductie van nieuwe passieve vormen

Nadat ik in Hoofdstuk 1 en 2 de achtergrond bij dit onderwerp en de onderzoeksvragen heb uitgelegd, zoals hierboven beschreven, onderzoek ik in **Hoofdstuk 3** de introductie van de drie nieuwe passieve vormen. Het doel van dit hoofdstuk is om vast te stellen wanneer deze passieve vormen precies zijn geïntroduceerd en ook wat de reden is voor deze introductie. Het is hierbij vooral van belang om vast te stellen of dit voor of na het verval van V2 is, en of er een grammaticale relatie is tot het V2-systeem. Ik laat in dit hoofdstuk zien dat de introductie van de drie speciale passieve vormen—de *prepositional passive*, de *recipient passive* en de *ECM passive*—nog dichter bij elkaar ligt dan eerder aangenomen in de literatuur, namelijk in de tweede helft van de veertiende eeuw. Dit betekent dus dat alle drie de vormen al in gebruik waren voordat het verval van

V2 begint, hoewel ze nog niet véél gebruikt worden en het gebruik nog niet helemaal regelmatig is. De prepositionele vorm is de eerste nieuwe passieve vorm die wordt gevonden in de manuscripten, maar ik beargumenteer dat de hele vroege voorbeelden die eerder in de wetenschappelijke literatuur zijn aangedragen niet onomstotelijk bewijzen dat de prepositionele passief al een geaccepteerde constructie is in de taal. Het corpus dat ik gebruik bevat geen eerdere voorbeelden van de *recipient passive* en *ECM passive* en de introductie blijft dus staan in de veertiende eeuw.

Een andere belangrijke conclusie die ik trek in hoofdstuk 3 is dat de nieuwe passieve vormen allemaal geïntroduceerd worden vanwege andere kleine veranderingen in de taal en dus niet vanwege een algemene ontwikkeling van passieve vormen in de taal. De reden voor de introductie van de *prepositional passive* is, beweer ik, het feit dat de prepositionele passief eerst op een onopvallende manier – voornamelijk in betrekkelijke bijzinnen waar het zogenaamde ‘achterlaten’ van het voorzetsel aan het einde van de zin al mogelijk was. De introductie van de *recipient passive* is toe te schrijven aan het verdwijnen van de vooropplaatsing van lijdende en meewerkende voorwerpen. Deze vooropplaatsing bereikt hetzelfde effect als de passieve vorm en zolang deze vooropplaatsing mogelijk is, is er geen reden om een *recipient passive* te gebruiken. De *ECM passive* wordt geïntroduceerd samen met de actieve versie, vaak toegeschreven aan invloed van het Latijn. Ik beargumenteer dat de passieve versie veel lijkt op vormen die al aanwezig zijn in het Engels en dus niet heel erg opvalt als nieuwe constructie. Dit is een groot verschil met de actieve versie, die opvalt als nieuwe constructie en waarschijnlijk daarom ook relatief weinig gebruikt wordt.

Interactie tussen syntaxis en informatiestructuur in het Oudengels

In **Hoofdstuk 4** onderzoek ik de interactie tussen syntaxis en informatiestructuur in het Oudengels, specifiek in het V2-systeem zoals dat in het Oudengels in gebruik was. Ik richt me hierbij op drie soorten zinnen: passieve zinnen, zinnen die beginnen met een voorop geplaatst lijdend voorwerp, en zinnen die beginnen met een voorzetselgroep (vaak bepalingen van tijd of plaats). De studies in dit hoofdstuk laten zien dat deze verschillende zinstypes allemaal beschikbaar zijn om de over het algemeen gewenste *given-before-new* volgorde van informatie te bereiken. Daarnaast wordt de eerste positie in de zin gebruikt door elementen van verschillende syntactische status—onderwerp, lijdend voorwerp, voorzetselgroep—en deze bevatten in de meerderheid oude informatie. De niet-onderwerpen in eerste positie bestaan voornamelijk uit informatie die eerder genoemd is, of te herleiden uit de context, en daarmee zijn ze dus *unmarked theme*. Ze vervullen daarmee een belangrijke functie in het verbinden van twee zinnen, door op de informatie uit de vorige zin door te gaan.

Ik laat zien dat in het Oudengels de lijdende vorm ook al wordt gebruikt om een *given-before-new* volgorde van informatie te creëren in de zin. De lijdende vorm in het Oudengels is nog wel minder gestandaardiseerd. Vooropplaatsing van lijdend

voorwerpen wordt in het Oudengels—in tegenstelling tot het hedendaags Engels—ook gebruikt voor *given-before-new*, maar het percentage is lager dan dat voor de passieve zinnen. Daarnaast valt op dat de uitzonderingen vooral contrastieve objecten bevat (tegenstellingen). Tot slot laat ik in hoofdstuk 4 zien dat voorzetselgroepen in eerste positie in de zin ook vooral oude informatie bevatten en gebruikt worden om een verbinding aan te brengen met de vorige zin. Maar ze kunnen ook weer contrastief zijn. Dit is anders dan de passieve zinnen, maar vergelijkbaar met de voorop geplaatste lijdend voorwerpen.

Samenvattend worden deze drie constructies allemaal gebruikt door sprekers om bepaalde informatie zo optimaal mogelijk te ordenen. Dat betekent dat (tot op zekere hoogte) deze drie syntactische mogelijkheden gelijkwaardige keuzes zijn voor de spreker voor het optimaal ordenen van de informatie. Dat de opties gelijkwaardig zijn is het duidelijkst voor zinnen met een voorop geplaatst lijdend en passieve zinnen omdat daarbij dezelfde elementen worden omgedraaid, het onderwerp en het lijdend voorwerp (vergelijk *Hem zag ik* en *Hij werd gezien door mij*).

Veranderingen in de interactie tussen syntaxis en informatiestructuur bij het verval van V2

In **Hoofdstuk 5** onderzoek ik het verval van V2 in de periode na 1500, in het bijzonder het verloop van de ontwikkelingen die beschreven zijn in hoofdstuk 4. De studies in het hoofdstuk laten zien dat het verval van V2 niet alleen bestaat uit een verandering in de positie van het onderwerp en de persoonsvorm ten opzichte van elkaar, maar dat vooral de functie en het gebruik van de eerste positie verandert door de eeuwen heen.

Het verval van V2 is al door veel taalkundigen onderzocht en de algemene ontwikkeling is inmiddels duidelijk. In het eerste deel van hoofdstuk 5 pluis ik een aantal aspecten van het verval van V2 verder uit. Ik kijk in dit deel van het hoofdstuk specifiek naar zinnen die beginnen met voorzetselgroepen omdat dit een voorbeeld is van een zin die variatie laat zien in de volgorde van onderwerp en persoonsvorm. Deze categorie van zinnen is nog minder beschreven in de literatuur. De studies laten zien dat gedurende de vroegmoderne periode (1500-1710) inversie bijna helemaal verdwijnt, met uitzondering van passieve zinnen en zinnen met een zogenaamd *unaccusative* werkwoord. Dit zijn werkwoorden waarbij het onderwerp niet echt de 'uitvoerder' is van het werkwoord, maar eigenlijk zelf de actie ondergaat, bijvoorbeeld in de zin *De mussen vielen van het dak*.⁹⁸ Vanaf ongeveer 1700, het begin van de Modern Engelse periode, komt inversie alleen nog voor zoals het ook nog in het hedendaags Engels in gebruik is—in hele specifieke contexten, vooral na plaatsbepalingen (*Into the room came an elephant*). Daarnaast laten de studies zien dat het Engels zich langzaam ontwikkelt tot een taal waarin hoofdzinnen

⁹⁸ In het Nederlands vormen deze werkwoorden in de regel een voltooid verleden tijd met *zijn*, en niet met *hebben*, zoals andere werkwoorden. Vergelijk bijvoorbeeld *De mussen zijn gevallen* met *De mussen hebben gevlogen*.

in de meeste gevallen (ongeveer 75%) beginnen met een onderwerp. Dit betekent dat de eerste positie in de hoofdzin veel minder gebruikt wordt door andere elementen (zoals een lijdend voorwerp of bepalingen van tijd en plaats). Het effect hiervan is dat als deze positie wel gebruikt wordt, dit redelijk bijzonder is en dit dan gebeurt vanwege een speciale reden, zoals het aangeven van een tegenstelling of nadruk. Vooral voor het voorop plaatsen van lijdend voorwerpen is dit nadrukkelijk het geval. Het onderwerp blijft dus tijdens de moderne periode over als de enige manier om de zin te beginnen met een *unmarked theme* dat bestaat uit oude informatie die niet contrastief is of waar geen nadruk op ligt. Dit is geen echte syntactische verandering, maar wel een verandering die heel bepalend is voor het Engels als taal, ook weer in vergelijking met de nauwverwante talen.

Gevolgen van het verval van V2

In **Hoofdstuk 6** kijk ik tot slot naar de gevolgen van de veranderingen in het gebruik van de eerste posities in de zin. Ik onderzoek een aantal zinstypes waarvan ik beargumenteer dat de toename in hun gebruik toe te schrijven is aan het toegenomen belang van het onderwerp als *unmarked theme*, omdat deze zinstypes er allemaal voor zorgen dat een element dat oorspronkelijk niet het onderwerp van de zin is, dit toch wordt. Als eerste gaat het dan om de passieve zinnen. Ik laat zien dat, zoals Seoanes (1999a) kleinere studie al suggereerde, het gebruik van passieve zinnen toeneemt in de vroegmoderne en moderne periodes. De drie nieuwe passieve vormen die waren geïntroduceerd in de veertiende eeuw worden langzaam maar regelmatig in hun gebruik en worden vaker gebruikt, net als alle passieve vormen. Tot slot kijk ik naar een aantal recentere ontwikkelingen, die er op neerkomen dat het onderwerp in het Engels vaker bestaat uit een element dat niet de traditionele rol van agens heeft ('uitvoerder'), zoals in de volgende zinnen.

- (9) The book sells well.
'Het boek verkoopt goed'

- (10) This tent sleeps four.
Deze tent slaapt vier [mensen]
'Er kunnen vier mensen in deze tent slapen'

In (9) is *the book* niet de agens van *sell*; het wordt zelf verkocht, maar verkoopt niet iets. In (10) is *This tent* niet de agens van *sleep*; het is niet de tent die slaapt. Hoewel deze constructies niet makkelijk te vinden zijn in de beschikbare corpora is er genoeg bewijs om aan te nemen dat hun gebruik toeneemt vanaf ongeveer de zestiende eeuw. Het is dus aannemelijk dat de ontwikkelingen na het verval van V2 hebben geleid tot een toename in het gebruik van dit soort constructies.

Conclusie

De studies in Hoofdstuk 3 tot 6 laten een aantal dingen zien met betrekking tot de vraag of syntaxis en informatiestructuur los van elkaar opereren of elkaar beïnvloeden. In het geval van het verval van V2 en de passieve vorm in de geschiedenis van het Engels lijkt het niet zo te zijn dat de behoefte van sprekers om aan bepaalde principes van informatiestructuur te voldoen daadwerkelijk leidt tot de introductie van nieuwe constructies in de taal. De constructies die ik onderzoek—de nieuwe passieve vormen, de non-agens onderwerpen, de passieve vorm *an sich*—zijn allemaal al aanwezig in de taal voordat het verval van V2 wordt ingezet, of worden om afzonderlijke redenen geïntroduceerd. Tegelijkertijd is het wel zo dat de principes van informatiestructuur bepalend zijn voor welke constructies veel worden gebruikt. Ten eerste wordt in het Oudengels de positie van het onderwerp grotendeels bepaald door de informatiewaarde van dit onderwerp. Ten tweede zijn er in het Oudengels verschillende zinstypes die een *given-before-new* volgorde van informatie kunnen creëren door de argumenten in de zin te herordenen. Wanneer enkele van deze opties uit gebruik raken (bijvoorbeeld vooropplaatsing van lijdend voorwerpen), neemt het gebruik van de overgebleven opties, voornamelijk de passieve vorm, toe. Dit proefschrift laat vooral zien dat in deze twee ontwikkelingen in de geschiedenis van het Engels zowel de syntactische veranderingen als de ontwikkelingen in het uitdrukken van informatiestructuurprincipes belangrijk zijn voor het begrijpen van wat er gebeurt in de taal. Samengenomen bieden de syntactische ontwikkelingen en informatiestructuurprincipes ons een inzicht in hoe het Engels zich heeft ontwikkeld, specifiek in de veranderingen die het Engels onafhankelijk heeft doorgemaakt ten opzichte van nauwverwante talen als Nederlands en Duits.

Curriculum Vitae

Gea Dreschler was born in Ede (the Netherlands) on 11 September 1986 and grew up in Veenendaal. After attending secondary school in Amersfoort, she moved to Leiden to do a BA in English Language and Culture. In her second and third year, she combined this with a minor in Journalism. She completed her BA in 2007, with a thesis on the syntax of negation in Dutch Sign Language. She decided to further specialize in linguistics with a research master *Linguistics: Structure and Variation in the Languages of the World*, also at Leiden University. She wrote her thesis about optionality in Minimalist syntax and graduated towards the end of 2008. By that time, she had already started on her PhD project at the Radboud University Nijmegen, of which this thesis is the result. While doing her PhD, she also worked as a part-time writing coach at the *Academisch Schrijfcentrum Nijmegen*. In September 2012 she started working as a Lecturer in English Linguistics at the Department of Language, Literature and Communication at VU University Amsterdam. She is now an Assistant Professor at the same department.